

Information Technology and Business

CHAPTER

1

Syllabus Mapping	Unit
Concepts of data, information and computer based-information system, impact of information technology on business [business data processing, intra-organizational and inter-organizational communication by using network technology, business process outsourcing and knowledge process outsourcing], types of information system- transaction processing system (TPS), management information system (MIS), decision support system (DSS), knowledge management system (KMS) and their implementation at managerial levels [operational, tactical and strategic].	Module I Unit I

1.1 CONCEPTS OF DATA, INFORMATION, AND KNOWLEDGE

The two words *data* and *information* are often used interchangeably but are actually not the same. While data is a collection of raw facts or figures, information comprises processed data to provide answers to the *who*, *what*, *where*, and *when* type of questions. There is one more important term that we must learn before going into the details, namely *knowledge*. Knowledge is the application of data and information to answer the *how* part of the question. The relationship among data, information, and knowledge can be clearly understood from Figure 1.1.

1.1.1 Features of Data

The features of data are listed here:

- Consists of raw facts or figures.
- Consists of unorganized and unprocessed facts.
- May be a simple statement.
- Can exist in any form, usable or not.
- Does not have any significance unless processed to derive information.
- Is obtained from experiments or surveys.
- Is used as a basis for making calculations or drawing conclusions.
- Can exist as numbers, text, images, sounds, and so on.
- An organization, sometimes, has to decide on the nature and volume of data that is required for deriving information. For example, date of birth of a person, salary of a customer, price of an item, and so on.



In the business world, data usually exists in the form of spreadsheets.

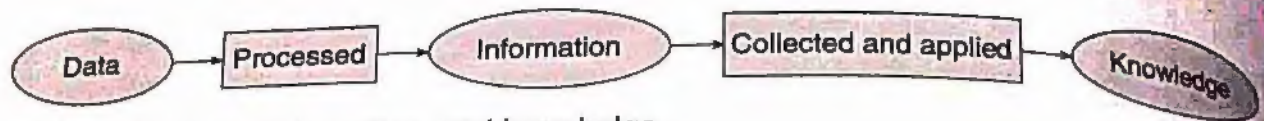


Figure 1.1 Data, information, and knowledge

1.1.2 Features of Information

The features of information are as follows:

- Consists of processed data.
- May be the collected facts and data about a particular subject.
- May be collected data, organized and presented in a systematic fashion, to understand the underlying meaning. For example, a telephone service may supply telephone numbers to a loan finance company. Here, the collected data (numbers), when collected and organized to be used by the company, becomes information.
- Embodies the understanding of certain relationships, possibly cause and effect. For example, whether a person with a given date of birth is eligible to vote, whether a customer with a specified salary is able to afford the product, and so on.

1.1.3 Features of Knowledge

The features of knowledge are as follows:

- Comprises human understanding of a subject matter that has been acquired through proper study and experience.
- Is based on learning, thinking, and proper understanding of the problem area.
- Makes the information more usable.
- Is a deterministic process.
- Represents a pattern that connects the pieces of information together.
- Provides a high level of predictability to explain what happened in the past and what will happen next.
- Comprises all the data and information learned through experience or study.
- Helps people draw meaningful conclusions. For example, how to market the product, so that it sells more.

Figure 1.2 shows the data, information, and knowledge (DIK) pyramid, in which data forms the basis of information and knowledge.

Example 1.1 If we know that Kellogg's cornflakes are too costly to be bought by a daily wage, a small packet of ₹10 can be introduced in the market, which he can purchase. Here, data comprises the price of the cornflakes and the wages of the labourer. Information is that he cannot afford to buy branded cornflakes. Knowledge is to reduce the price and weight, to make it affordable for him.

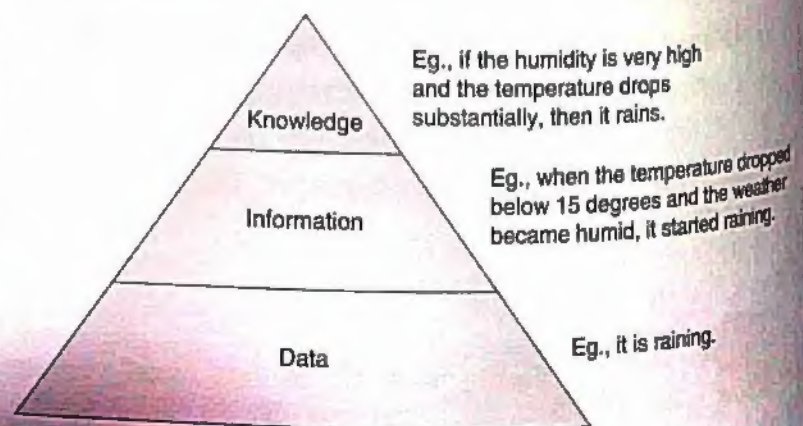


Figure 1.2 The DIK pyramid



Knowledge is not information and information is not data. Knowledge is derived from information, which, in turn, is derived from data.

1.1.4 Characteristics of Information

To be useful, information must possess certain characteristics, as shown in Table 1.1.

Table 1.1 Characteristics of information

Characteristics	Utility	Examples of unusable information
Relevance	Information must pertain to the problem at hand. It should be applicable in that context.	The monthly income of a candidate's father is irrelevant for appearance in the CA examination.
Complete	Partial information is worse than no information.	To market a product, the marketing team must have information about customer buying patterns as well as their incomes. If a product is promoted to high-income group customers, who usually do not prefer to buy it, it often leads to wastage of time and effort.
Accurate	Erroneous information may result in wrong decisions.	If the consumer survey gives wrong information about users' preferences, then the decisions taken, in light of the survey results, will be always wrong.
Current	The information used must be recent and the latest.	A person cannot buy a stock today, referring to the price and profit margin of that stock on a previous date.
Economical	The cost of obtaining information should not be so high that it cannot be justified.	If the cost of conducting a survey about a new product is too high, then it may diminish the profit from sales.

1.2 IMPACT OF INFORMATION TECHNOLOGY ON BUSINESS

Businesses have been at the forefront of technology for ages. Whatever can speed up the production process will draw in more business. With the widespread use of computers, a new age of information technology (IT) has begun. To reap its benefits, businesses are rapidly changing their infrastructure. These days, almost every business is conducted over computers and communication devices, in order to organize large databases, personal schedules, and various other forms of essential information.

Information technology is a comprehensive term that includes all types of technology used to exchange, store, use, or create information. Commonly used devices to support IT include computers, servers, peripheral devices, Internet connectivity equipment, and phone systems. Breakthrough in IT has resulted in better or automated solutions that have not only increased the productivity, but also

involve lower operational costs, improved speed, ease of sharing and storing information, decrease in human error through automation, and increase in revenue.

IT is rapidly permeating the business world, affecting how companies make and market their products, as well as how people communicate and accomplish their jobs. Specialized software shapes best practices and industry standards, continually changing the face of business in almost every way. Let us see how the use of IT has revolutionized the way we do our business.

1.2.1 Business Data Processing

Business data processing (BDP) means performing operations in order to convert business data into useful information. IT enables businesses to store large amounts of information, summarize them, and analyse them, in order to understand the business in a better way. Once data is entered into a spreadsheet application, users can draw a variety of charts to interpret the data from different angles. For example, view the sales of a product in a particular region during the entire year, view the sales of a product in all the regions in a particular month, and the like.

Another example could be from banking. Banks store records of millions of users and keep complete information about their accounts, loans, fixed deposits, and so on. They keep sending statements and other promotional schemes to their customers from time to time.

Other areas in which IT has a major impact on BDP include:

Accounting and Payroll

Businesses can use software accounting packages, like Microsoft Money or QuickBooks to manage their accounts. No longer do they need to hire full-time or third-party accountants. These software packages enable business managers to generate informative reports and financial statements, in order to make effective business decisions. Accounting systems keep track of every rupee a company spends, along with every rupee of revenue.

The accounting software also stores payroll information, tax records, and other specialized data. Business without these software packages cannot be even imagined today. Besides increasing efficiency, IT also reduces time and space required to store a large number of files with hundreds of papers in it. Can you think of searching information from these files?

Inventory Management

Businesses can use software packages like Microsoft Excel or Open Office to store data, pertaining to inventory, sales, and receivables and payables. The inventory management system keeps track of products in the inventory and ensures that the business neither runs out of stock nor is in excess.

Enhanced Productivity

Technology tools help employees do their tasks quickly and efficiently in very less time, and that too without any error. Therefore, empowering employees with the right hardware and updated software keeps them working at their best.

Easier Storage

Information technology eliminates the need for double or triple entry systems and reduces paperwork to a considerably large extent. Details about contracts, quotations, customers, suppliers, dealers, retailers and so on can all be stored in databases and accessed in minutes.

Automated

With the use of IT, businesses can hire and keep track of their employees like a machine. It can cut down on the cost of maintaining a low cost

Economic

Computers can be used to perform one of the functions of a lower level of automation.

1.2.2

With the use of IT, businesses can hire and keep track of their employees like a machine. It can cut down on the cost of maintaining a low cost

Inter

Today, the Internet has become a part of our daily lives. It has changed the way we do business. It has made it possible for businesses to reach a wider audience. It has also made it possible for businesses to interact with their customers in a more personal way.

With the use of IT, businesses can hire and keep track of their employees like a machine. It can cut down on the cost of maintaining a low cost

With the use of IT, businesses can hire and keep track of their employees like a machine. It can cut down on the cost of maintaining a low cost

With the use of IT, businesses can hire and keep track of their employees like a machine. It can cut down on the cost of maintaining a low cost

With the use of IT, businesses can hire and keep track of their employees like a machine. It can cut down on the cost of maintaining a low cost

With the use of IT, businesses can hire and keep track of their employees like a machine. It can cut down on the cost of maintaining a low cost

With the use of IT, businesses can hire and keep track of their employees like a machine. It can cut down on the cost of maintaining a low cost

With the use of IT, businesses can hire and keep track of their employees like a machine. It can cut down on the cost of maintaining a low cost

With the use of IT, businesses can hire and keep track of their employees like a machine. It can cut down on the cost of maintaining a low cost

Automation

With technology, businesses can automate certain functions that have historically required the need to hire an employee. They can use software like Quicken and Quickbooks to perform frequent book-keeping functions. Similarly, the sales function can be automated through contact management sites like Salesforce. Using automated software enables business managers to concentrate on strategy and cut down on labour expenses. Automation technology, such as assembly lines and computer-controlled machinery, enables industries and organizations to produce large volumes of goods in less time and at low cost.

Economic Efficiencies

Companies can harness IT resources to lower their costs. Using IT infrastructure, redundant tasks can be centralized at one location. For example, a large company could centralize their payroll function at one location to lower employee costs. Economic efficiencies can also be realized by migrating high-cost functions into an online environment. Companies can offer customers email support, which may have a lower cost than a live customer support call. Cost savings can also be made through outsourcing opportunities, remote work options, and lower-cost communication options.

1.2.2 Enhanced Intra-organizational and Inter-organizational Communication

With IT, an organization can be benefitted from faster, efficient, effective, and low-cost communication systems, as explained in this section.

Internet Marketing

Today, all businesses are moving towards making their presence felt on the Internet. Whether big or small, all businesses have their websites to advertise their products, take orders, buy merchandise, and sell excess products. Many businesses such as Flipkart, Amazon, Snapdeal, OLX, and so on, operate only online. Some businesses are also doing Internet marketing on other social websites. For example, we all have seen advertisements of products on Facebook.

With enhanced communication (Figure 1.3), companies are able to build strong relationships within their value chains. As a result, with extensive use of technology, even a small business can compete with established and big business houses. Using advanced video and graphics editing software, business owners create professional marketing materials that reach out to millions of netizens worldwide. They use cost-efficient web marketing tools like Google AdWords to spread targeted marketing messages, either to a broad audience or to a selected group. Business owners also share expensive advertising space online through banner- and traffic-swaps.

Additionally, through email marketing, business managers are effectively reaching out to a large group of people with newsletters, coupons, or other means of business updates, and that too at very low costs. Going a step ahead, they have also started using mobile marketing to offer discounts to customers, through text messaging and by advertising their products and services. Undoubtedly, technology has freed small businesses from the restrictions of print advertisements to reach out to their customers. IT has helped businesses minimize geographic communication barriers.



Figure 1.3 Enhanced communication

Technology has brought businesses closer to customers. Businesses use emails to reply to customers' queries, get their feedback, offer online chat to help them navigate through their websites, and use call centers to support their customers every moment.

Stay Competitive

To survive in the marketplace, every business is racing to get the technology tag. Using Internet and online surveys, business managers stay up-to-date with recent trends, techniques, software, and human resources. Managers communicate with their employees in different locations, resolve both employee and customer issues in almost no time, and keep track of their status as well as competitors' positions through different postings on the web.

With better communication, business managers can manage their time better, freeing themselves from administrative minutiae and focusing on more productive activities, such as making vital contacts with suppliers and potential strategic partners. Business managers and entrepreneurs can hold face-to-face meetings online, from anywhere in the world. For example, IndusInd bank has started mobile face-to-face banking.



Companies with first mover strategy use IT to create new products that give them a cutting edge.

Improved Sharing of Information

With technology, any type of information—written or broadcast—can be shared more quickly and with fewer resources. Besides sharing information about new products with customers, IT also helps send inventory orders, make payments, and invite quotations online. Information technology also enables companies to get into e-learning and other forms of online training to train their staff in a cost-effective manner.

Customer Relationship Management

A customer relationship management (CRM) system is used to track a customer throughout his experience or interaction with the business. Managers can view the information about how many times the customer has called to order a product or service, for help, with a technical question, or for any other conversation. All these details help to build strong relationships with customers, by providing them with customized and proactive responses.

Telecommuting, Teleconferencing, and Videoconferencing

Business houses are using telecommuting to help employees stay in touch with each other from different locations. Employees use email, online collaboration tools, and mobile computing devices to share digital documents, convey information through presentations, and create training videos for new joiners.

Teleconferencing over the phone and video brings together web cameras and audio and collaborative online meeting spaces to create an interactive environment, in which the participants can see each other, work together on documents, and recreate the in-person meeting experience, no matter where they are in the world. This enables businesses to reach global customers and partners.



Nowadays, a business that is not conducted online or one that has an outdated and unattractive-looking website is considered to be unprofessional.

1.2.3

The term even res process these tw

From

Figure

focus that is outso count BPO try is Th Figur back sour billi in fr such as sh



Adv

Ri are

1.2.3 Outsourcing

The term *outsourcing* implies hiring employees who work outside the company. These employees may even reside in some other country. These days, we frequently hear that people are working in business process outsourcing (BPO) or knowledge process outsourcing (KPO). In this section, we will discuss these two forms of outsourcing and understand the underlying difference between the two.

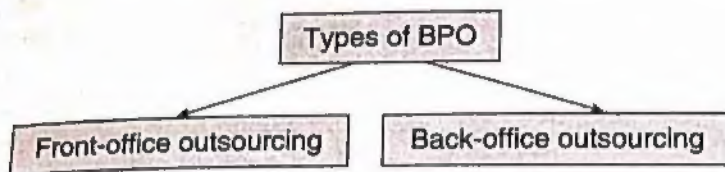


Figure 1.4 Types of BPO

focusing on routine business tasks. The BPO service that is outsourced to another country is called *offshore outsourcing*, while the one outsourced to a neighboring country is called *nearshore outsourcing*. Similarly, the BPO service that is outsourced within the same country is called *onshore outsourcing*.

There are two main types of BPOs (as shown in Figure 1.4), which are front-office outsourcing and back-office outsourcing. While in back-office outsourcing, internal business functions such as payroll, billing, purchasing, accounting, and HR are outsourced, in front-office outsourcing, customer-related services such as marketing and technical support are outsourced, as shown in Figure 1.5.



Call centre jobs are BPO jobs.

Business Process Outsourcing

Business process outsourcing is the practice in which employees outside the company are hired to perform specific business tasks like payroll. Outsourcing is done as a cost-saving measure by big companies that want to focus on their key business area, rather than



Figure 1.5 Services provided by BPO

Advantages The main advantages of BPOs are as follows:

- Companies can focus on their specialized areas and specific competencies.
- Speed and efficiency of routine business tasks enhance when these jobs are outsourced to a third party.
- All the employees of the company can concentrate in designing and implementing the core business strategies to bolster competitive advantage and enhance value chain engagement.
- Companies in developed nations usually outsource their routine tasks to developing nations, where they have to pay very little for the same job, thereby helping to cut involved costs.

Risks There are, however, certain serious risks that exist while outsourcing business processes. These are as follows:

- Data privacy breaches can take place.
- There can be underestimated running costs.
- There can be overdependence on service providers.

Knowledge Process Outsourcing

Knowledge process outsourcing is another form of outsourcing, in which a company outsources core business activities competitively important or form an integral part of its value chain. It requires advanced analytical and technical skills, as well as a high degree of specialist expertise. The KPO services include all kinds of research and information gathering activities like intellectual property research for patent applications, equity research, business and market research, legal and medical services, training, consultancy, and research and development in fields such as pharmaceuticals, biotechnology, animation and design, web development, CAD/CAM applications, fraud analytics, project management, remote education, research and development, radiology, and medical transcript preparation, as shown in Figures 1.6 and 1.7.



Figure 1.6 Services provided by KPO

KPO sector revenue by segment: 2010

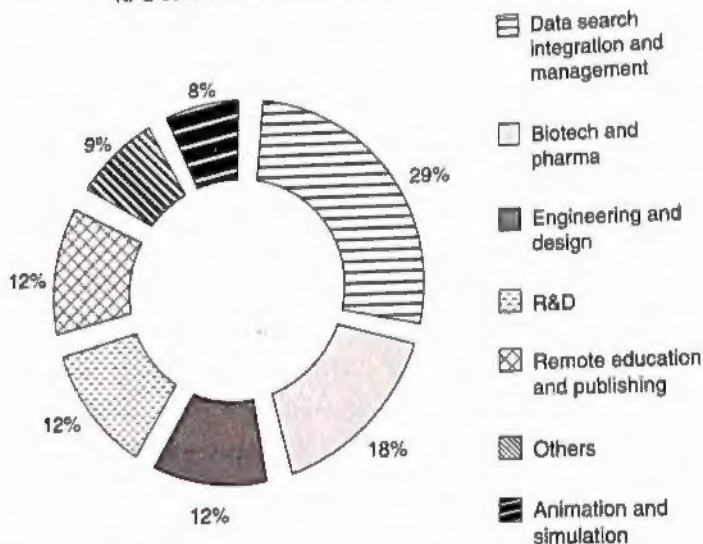


Figure 1.7 KPO sector services

The KPOs basically exist not only as a measure of cost reduction, but also as a compensation for shortage of skilled labour. Developed nations such as UK, US, Canada, and England see developing countries such as India, Sri Lanka, Chile, Mexico and Eastern Europe, especially Poland, Romania, and the Baltic States, as suitable regions to set up a KPO.

A KPO extends the work of a BPO by adding more business complexities. It is always recommended that only a company that has a good knowledge of running a BPO must get into the KPO market.

Advantages The main advantages of KPOs are as follows:

- They reduce cost.
- They compensate for shortage of skilled employees.
- They offer low cost expertise.
- They offer employment for educated people in developing countries and benefits of their economies.
- They add value to the company's core competencies.

Risks There are, however, certain risks that involve in KPOs. These are as follows:

- Important data of the company is passed to a third party, which may lead to security breaches.
- The character and quality of the work of employees in another company, and may be in another country, cannot be assured.

- Establishing a KPO takes time, and thus, cannot provide immediate solutions to prevailing problems.
- There can be increased complexities, due to lack of communication, between the company and the KPO, possibly because of legal, language, and cultural barriers.
- Inferior quality of outputs, missed delivery times, and cultural differences can take place.
- Increased costs due to salary inflation, increased standard of living, and improved (stronger) currency of some developing countries like India may adversely affect the company.
- Political and economic instability as well as terrorist activities in the countries can affect the company.
- Losing out on key talent at home is a major threat to the company.



KPOs provide high-end services to customers and require skilled personnel. Additionally, the salary of employees working in KPOs is far better than those working in BPOs.

Reasons for successful KPOs in India Following are some of the reasons for the success of KPOs in India

- Maturity in handling BPOs
- Availability of high-end skilled labour, including chartered accountants, engineers, doctors, architects, lawyers, economists, biotechnologists, and the like
- Low wages and operational costs as compared to developed nations
- Operational efficiency
- Adherence to international quality standards

These days, however, many companies are moving their voice-based KPOs to countries (such as Chile and Mexico) with higher political stability and similar language, time zone, and cultural alignments. People in these countries have English as their first language and use the same accent as that used in most of the core companies.

Example 1.2 Delta and United Airlines brought back their outsourced customer service operations to North America, when low rates of satisfaction for more sophisticated conversations with customers were reported. Even the other employees in the core company, who were dealing with outsourced processes, had the same experience.



While basic education is necessary to get a BPO job, for a KPO job, competency in a particular field is a must. A job in a KPO is not suited for a person who is a *Jack of all trades*. In a KPO, employees are rigorously trained in a specific sector.

These days, BPOs are moving towards Philippines and Bangladesh. In such a situation, India is concentrating more on KPOs. US stands strict on outsourcing BPO, but it can never curb outsourcing of activities that require skilled labour, as there is an enormous deficiency of such labour force in the country.

1.3 TYPES OF INFORMATION SYSTEMS

Business organizations today leverage available technologies to manage their business, collect data, process data to analyse it to make important decisions, interact with customers and partners, curtail costs, and generate revenues. For performing these operations, business organizations use different types of information systems, wherein an *information system* is a collection of hardware, software, data, people, and procedures that are designed to generate information to support routine activities.

There are different types of Information systems. But before reading about them, let us first understand why we cannot have a single information system deployed in an organization.

1.3.1 Why are There Different Types of Information Systems?

Earlier, when computers were new to the world of business, each time an information system was needed, it was *tailor-made* or customized for the specific purpose. Computer experts soon realized, however, that most of the information systems were designed to solve a similar set of problems. While attempts were made to build a single system that would solve a whole range of similar problems. While designing information systems for similar problems, it was important to define how, where, why, and by whom the information system would be used. These questions gave rise to different information systems that were designed for different types of users working in the organization.

1.3.2 Implementation of IS at Managerial Level (Operational, Tactical, and Strategic)

To identify different types of information systems, classification is done to identify a set of activities that can be categorized together, so that they can be treated as a single unit. Although there is no well-defined rule to classify an information system, it is necessary to classify them in order to ease the design of applications they would render. Figure 1.8 shows the pyramid model used to classify information systems. As most organizations work in a hierarchy, the information systems are also categorized to follow the hierarchy.

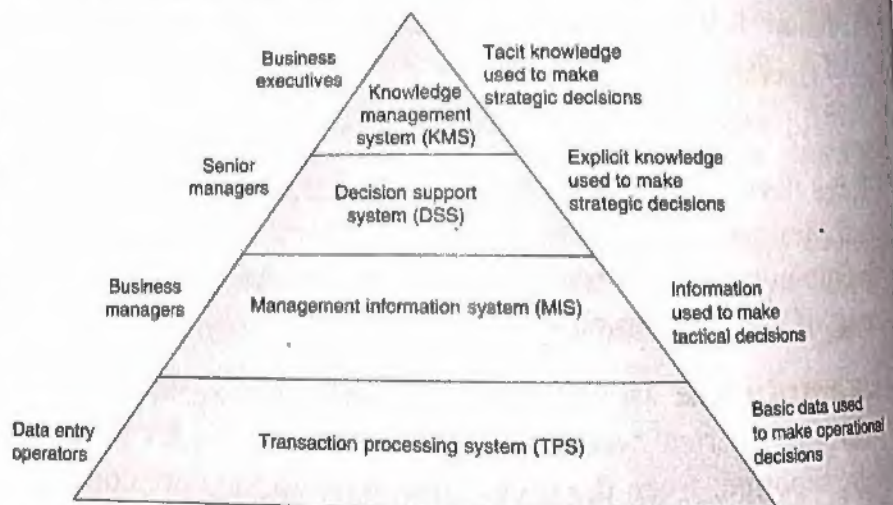


Figure 1.8 Pyramid model of information systems

Figure 1.8 gives a generalized pyramid model for classifying different types of information systems in an organization. It does not mean, however, that every business organization must strictly adhere to it. While some small businesses can have just three levels, that is, transaction processing system (TPS), management information system (MIS), and decision support system (DSS), large organizations can have two additional information systems, namely office automation/information system (OAS/OIS) and executive information system (EIS). In this section, we will read about all of them.

Office Information System or Office Automation System

Office information system (OIS) or office automation system (OAS) uses hardware, software, and networks to enhance work flow and facilitate communication among employees. While hardware includes computers equipped with modems, video cameras, speakers, microphones, scanners, and fax machines, software comprises word processing, spreadsheets, databases, presentation graphics, email, web browsers, web page authoring, personal information management, and groupware to support the aforementioned activities. Besides hardware and software, an OAS also uses technology for communication, such as voice mail, fax, videoconferencing, and electronic data interchange (EDI) for the electronic exchange of documents, which include text, graphics, audio, and video (refer Figure 1.9).

In such a system, employees perform their tasks electronically rather than manually. For example, in an OAS in your college, the time table, academic calendar, subjects in the semester, along with their guidelines, notification of events in the university, fee structure, and the like, are all published online, either on the website or as notifications, using emails. Had it been a manual system, the office staff would have to take a photocopy of the information and send it to each student's home as a letter or as registered post.

Office automation systems are implemented to support a range of business office activities like creating and distributing graphics and/or documents, sending messages, scheduling, and accounting. Although it is meant to be used by clerks and officers working at the lower level of the pyramid, all levels of users avail some benefit from the OAS.

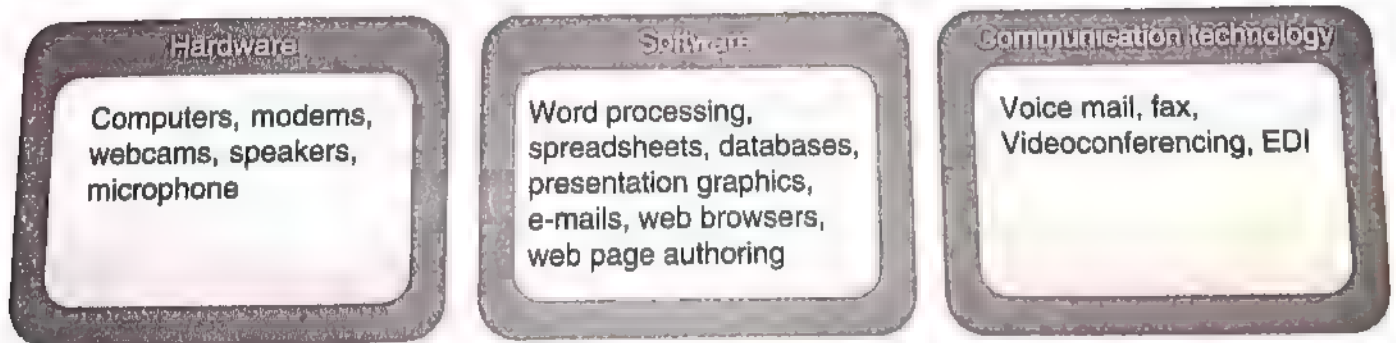


Figure 1.9 Components of OAS

Transaction Processing System

A transaction is a business activity (like deposit, payment, order, or reservation). Correspondingly, TPS is an information system that is specifically designed to capture and process data generated in an organization, during its day-to-day transactions. It is basically used by clerical staff, front-office personnel, and other employees working at the lowest level of the pyramid to perform the following transactions online.

- Record information regarding a student during his registration, an order placed by the customer, payment details of a client, and the like.
- Confirm an action or send a response to inquiries, issue receipts of payments, send a thank-you mail to customers, and so on.
- Collect customers' feedback.
- Generate employees' paycheck.
- Add, update, and delete existing data.

Transaction processing system was one of the first computerized systems developed to process business data (also known as *data processing*). The first TPS was based on batch processing, in which data is collected over a period of time, and all transactions are processed later as a group. But with drastic improvements in computing technology, batch processing systems were upgraded to online transaction processing systems (OLTPs), in which the computer processes transactions as they are entered.

These days, the OLTPs have widely replaced the earlier and now obsolete batch processing systems, and you have already seen OLTPs when registering yourself in college, shopping in a showroom that generates electronic invoices, paying bills in a restaurant that generates a printed receipt, and so on. Other places in which TPS is used include payroll systems, order processing systems, reservation systems, stock control systems, and systems for payments and funds transfers.

Advantages Here are some of the advantages of TPS:

- Enhanced efficiency and accuracy of business activities
- Faster processing
- Reduced clerical costs
- Improved customer service

Example 1.3 In an online air ticket booking system, travellers select their flight schedules and desired seats, and the system updates the availability of seats. The system also generates a bill and a copy of the ticket as its output.



Data gathered using a TPS can be used by upper level employees to view a detailed report and create an action report.

Management Information System

While TPS is best suited for routine transaction processing, business managers felt the need for an information system that could perform rapid calculations and data comparisons, in order to produce meaningful information for management. This need led to a new type of information system called MIS. It is an information system generating accurate, timely, and organized information that helps business managers make decisions, solve problems, supervise activities, and track progress by generating useful reports on a regular basis. Therefore, it is also known as a *management reporting system (MRS)*.

For better integration with business activities, MIS is integrated with TPS, so that MIS can generate reports using the data collected by TPS. For example, while TPS records the daily sales, updates the customer's account balance, and makes a deduction from inventory, MIS can use this information to produce reports that recap daily sales activities, list customers with past due account balances, identify slow or fast-selling products, and highlight inventory items that need reordering. It generates three types of reports, as shown in Figure 1.10. These are explained as follows.

Detailed report This lists transaction processing activities. For example, a detailed order report enlists all transactions that involve ordering of items.

Summary report This aggregates data so that managers can get a quick overview of the business activity. It synthesizes large amounts of information that contains totals, tables, or graphs. For example, an inventory summary report summarizes the items and the number of units available.

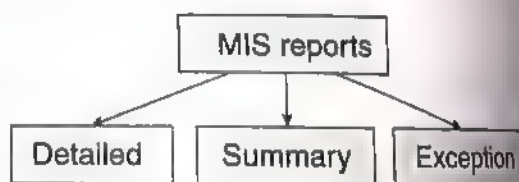


Figure 1.10 Types of reports

Exception reports are generated when an exception condition occurs. For example, an exception report might list all orders that are more than the specified number of days late. Managers use these reports to identify conditions.

Managers use these reports to identify conditions. Managers use these reports to identify conditions. Managers use these reports to identify conditions.

Managers use these reports to identify conditions. Managers use these reports to identify conditions. Managers use these reports to identify conditions.

Decision Support System

Transaction processing systems are used by users on a regular basis. They are used for term decisions. They are used for time, say, for example, for a decision to find a DSS.

In order to make a decision, managers also make use of the information on the Internet. They use the information of customers.

Decision support systems are tools that affect the decision-making process. Each person has more information.

Executive reports are used by executives. They are used for sales performance.

Executive reports are used by executives. They are used for sales performance. Executive reports are used by executives. They are used for sales performance.

Executive reports are used by executives. They are used for sales performance. Executive reports are used by executives. They are used for sales performance.

Executive reports are used by executives. They are used for sales performance.

Exception report This filters data to identify information that is beyond a normal condition (also called *exception condition*) and notify business managers, so that corrective measures can be instantly taken. For example, an inventory exception report notifies the purchasing department of items whose stock is less than the specified normal stock. These items must be reordered at the earliest. Exception reports enable managers to save their time, as they no longer have to go through the detailed report to identify exceptional conditions.

Management information system is used by business managers working at the middle level of the pyramid in order to ensure smooth running of the organization in the short- to medium-term. The highly structured information, produced by MIS, enables managers to evaluate their organization's performance, by comparing current output with the previous one.

Management information system supports structured decisions. Besides action and detailed reports, MIS also generates summary reports. Examples of MIS include sales management system, inventory control system, human resource management system, budgeting systems, and the like. The downside of these systems, however, is that they have little analytical capability. It also helps compare past values with the present, but does not facilitate prediction of the future.

Decision Support System

Transaction processing system and MIS are very well suited for providing information to appropriate users on a regular basis. But senior managers need some other type of information to make some longer term decisions. For short-term decisions, MIS is enough, but for decisions regarding a longer period of time, say a year, an MIS is not the one we should look for. For example, a sales manager needs information to find out how high to set yearly sales quotas, based on increased sales and lowered product costs. A DSS provides information to support such decisions.

In order to make vital decisions, DSS not only uses data from its internal information systems, but also makes use of data from external sources, like business magazines, surveys of competitors available on the Internet, interest rates, population trends, customer demographics, spending behavior of a group of customers, and so on.

Decision support system makes full use of query language (discussed in Chapter 5), statistical analysis tools, spreadsheets, and graphics to analyse data, evaluate results, and identify and document factors affecting a decision. This could not only help the sales manager to estimate the expected sales volume at each price level with greater accuracy, but also enable him to ask what-if questions, by changing one or more of the factors (that could affect the sales) and view the projected results.

Executive information system Executive information system is a special type of DSS, which is specifically designed for the information needs of business executives. It makes use of charts and tables to represent information that could help executives view trends, ratios, and other statistics, predict future sales patterns, summarize current costs, and forecast sales revenues.

Executive information system is basically used to make strategic decisions, and therefore, relies heavily on external sources of data like the Dow Jones news and the Internet, to retrieve information on interest rates, commodity prices, and other leading economic indicators. To store all the external as well as internal data, current as well as historical data, DSS and EIS use extremely large databases, called data warehouses (refer to Chapter 5).

Data warehouses, along with EIS, help executives analyse data according to the entire business or at the individual department, region, or a particular store. For example, the executive of McDonald's can view the daily sales, either in all the stores worldwide, or in a particular county, or in a particular city, or at a specific outlet.



Decision support system and executive information system are interactive systems that are used to analyse existing structured information to perform what-if analysis to generate summary reports and business forecasts to identify long-term plans and make strategic decisions.

Example 1.4 Financial institutions like banks use EIS to develop credit risk models to analyse the number and extent of lending or credit that is given to various sectors. For this, various techniques and formulae are used to determine the probability of loan defaults.

Knowledge Management System

In Section 1.1, we have already seen the difference between data, information, and knowledge. Knowledge is personalized information, a state of knowing and understanding that is stored and manipulated, to influence actions. Knowledge is obtained through the organization's data warehouse, DSS tools, internal expertise, and networks of knowledge workers.

All the discussed systems fall under KMS. We can also say that KMS extends the already existing systems by assimilating more information. Basically, KMS refers to a system that manages knowledge in an organization (in the form of documents, policies, procedures, expertise, and experience) to support identification, creation, capture, evaluation, storage retrieval, sharing, and dissemination of information, as shown in Figure 1.11. It helps employees have ready access to the organization's documented base of facts, sources of information, and solutions, as shown in Figure 1.11. It also provides network maps of the organization, depicting the flow of communication between entities and individuals.

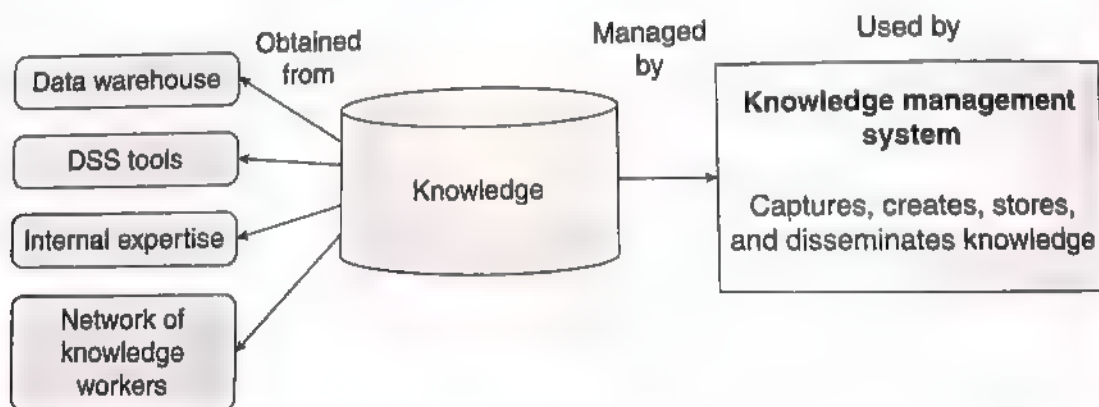


Figure 1.11 Knowledge management system

Example 1.5 The most effective strategy to promote a product in rural areas can be shared organization-wide, to help the sales and marketing team throughout the country to use the same strategy. It also opens up a new platform to discuss new and improved strategies.

Features

- Stores knowledge that is developed collectively.
- Comprises a range of practices to create and enable adoption of knowledge, embodied in an individual or in organizational processes.
- Knowledge exists in the form of documents, policies, procedures, expertise, and experience.
- Distribution of knowledge through KMS results in continuous change, reconstruction, and application of knowledge in different contexts.
- Required information is effectively and efficiently retrieved, whenever required.

Types of knowledge managed An effective KMS provides a variety of information, which includes the following.

Experiences It collects information retrieved from people, who have already tackled a problem and found a solution. Therefore, KMS stores information about successful strategies, and also the best practices to solve a particular problem.

Expertise location It stores a network map that informs the users about where to locate for an expert or expert information.

Communities of practice It enables groups of individuals to discuss their problems, opportunities, lessons learned, and other information gained from users.

Feedback An organization can collect feedback from customers and employees and share it with their research and development (R&D) departments, so that an integrated approach can be taken to understand the shared information and produce better products and services. For example, a company developing a new product may conduct research on their competitors and ask the R&D team to identify the ingredient needed in their product. The R&D team can find out from the KMS team the assets and processes the company has in place, which can be used to meet the sales potential, meet customer needs, and fill gaps within the marketplace.

Share project files An employee team can work collaboratively on a project. They can share their files and information to allow everyone in the team to upload and comment on the work performed by others.



Knowledge management system is widely implemented in firms of lawyers and for those working with accounting and management consultancy.

Advantages Knowledge management system can be associated with open source software, and open standards, open protocols and open knowledge licenses, and initiatives and policies. Its advantages include the following:

- Improved performance
- Competitive advantage
- Innovation, which gives a cutting edge to perform better than one's competitors
- Integrating and sharing of knowledge
- Solving problems faster
- Developing professional skills
- Orientation and training of new employees
- Support of better sales, by helping business users make critical decisions in a better and informed manner
- Creation of knowledge directories
- Avoidance of re-inventing the wheel, thereby reducing redundant work
- Retention of intellectual property, even after the employee leaves the organization
- Service as central repository, to retain information in a standard format
- Ensuring of consistent and speedy responses to user queries

Many large organizations are reaping the benefits of KMS. These days, even small businesses are implementing it.

Summary

- Information must be relevant, accurate, complete, current, and economical.
- These days, almost every business is conducted over computers and communication devices to organize large databases, personal schedules, and various other forms of essential information.
- Breakthrough in information technology (IT) has resulted in better or automated solutions that have not only increased the productivity, but also involve lower operational costs, improved speed, ease of sharing and storing information, decrease in human error through automation, and an increase in revenue.
- All businesses today are moving towards marking their presence on the Internet. Whether big or small, all business have a website to advertise their products, take orders, buy merchandise, and sell excess products.
- Outsourcing is done as a cost-saving measure by big companies that want to focus on their key

business area, rather than focusing on routine business tasks.

- An office information system (OIS) or office automation system (OAS) uses hardware, software, and networks to enhance work flow and facilitate communication among employees.
- MIS is an information system that generates accurate, timely, and organized information. This helps business managers make decisions, solve problems, supervise activities, and track progress, by generating useful reports on a regular basis.
- Decision support system (DSS) not only uses data from its internal information systems, but also makes use of data from external sources such as business magazines, surveys of competitors available on the Internet, interest rates, population trends, customer demographics, and spending behaviour of a group of customers.
- Knowledge exists in the form of documents, policies, procedures, expertise, and experience.

Glossary

Artificial intelligence (AI) A branch of computer science that works on making computers intelligent like humans

Back-office outsourcing Outsourcing of internal business functions such as payroll, billing, purchasing, accounting, and HR

Business data processing (BDP) Performing operations to convert business data into useful information

Data A collection of raw facts or figures

Expert system A branch of AI that designs intelligent machines, which solve real-world problems by using deductive logic

Front-office outsourcing Outsourcing of customer-related services such as marketing and technical support

Information Processed data to provide answers to *who, what, where, and when* type of questions

Information system A collection of hardware, software, data, people, and procedures that are designed to generate information to support routine activities

Information technology (IT) A comprehensive term that includes all types of technology used to exchange, store, use, or create information

Knowledge Application of data and information to answer the *how* part of the question

Knowledge management system (KMS) A system that manages knowledge in an organization (in the form of documents, policies, procedures, expertise, and experience) to support identification, creation, capture, evaluation, storage retrieval, sharing, and dissemination of information

Mobile application (app) Software application that runs on smartphones, tablets, and other mobile devices

Natural language processing (NLP) A branch of AI that involves programming computers to understand natural human languages

Nearshore
ing countr
Offshore
country
Onshore
country

Multi

1. Co
cal
(a)
(b)
2. Pr
(a)
(b)
3. T
(a)
(b)
4. I
(a)
(b)
- 5.

6.

7.

8.

Nearshore outsourcing Outsourcing to a neighboring country

Offshore outsourcing Outsourcing to another country

Onshore outsourcing Outsourcing in the same country

Outsourcing Hiring employees who work outside the company

Robotics A branch of AI that programs computers to see, hear, and react to sensory input

Multiple-choice Questions

- Collection of raw facts or figures is called _____.
(a) data (b) information (c) knowledge (d) output
- Processed data is known as _____.
(a) input (b) information (c) knowledge (d) output
- The 'how' part of a question is answered by _____.
(a) data (b) information (c) knowledge (d) output
- Information should be _____.
(a) economical (b) relevant (c) both of these (d) only (b)
- If the employee works hard and is skilled with the latest technology, then his promotion is guaranteed. Which of the following best describes this statement?
(a) data (b) information (c) knowledge (d) output
- Monthly salary of a parliament member is an example of _____.
(a) data (b) information (c) knowledge (d) output
- _____ means performing operations in order to convert business data into useful information.
(a) Business data processing (b) Built-in data process (c) Business digital processing (d) Business data management
- _____ software is widely used by businesses to analyse and summarize data and draw charts.
(a) Word processing (b) Spreadsheet (c) Database (d) Presentation
- BDP cannot be used for _____.
(a) accounting (b) payroll (c) inventory management (d) none of these
- Microsoft Money and QuickBooks are examples of _____ software.
(a) accounting and payroll (b) only payroll (c) inventory management (d) only accounting
- Inventory management can be done on _____ software.
(a) Microsoft Word (b) Microsoft Excel (c) Open Office Calc (d) (b) and (c)
- QuickBooks is a software used for _____.
(a) accounting (b) payroll (c) book keeping (d) all of these
- Flipkart, Snapdeal, Amazon are _____ websites.
(a) online shopping (b) social networking (c) email (d) search engine
- Online marketing is done through _____.
(a) email (b) Google AdWords (c) both (a) and (b) (d) none of these
- _____ was the bank which first started mobile face-to-face banking.
(a) ICICI (b) HDFC (c) Indus Ind (d) SBI
- CRM stands for _____.
(a) call record management (b) customer record management (c) customer relationship management (d) all of these

17. Web cameras and microphones are together used in _____.
 (a) tele conferencing
 (b) tele commuting
 (c) video conferencing
18. Outsourcing is mainly done to _____.
 (a) reduce costs
 (b) save time
 (c) increase productivity
 (d) automation
19. Payroll processing, if outsourced, will be a part of _____.
 (a) BPO
 (b) KPO
 (c) PPO
 (d) LPO
20. BPO service that is outsourced within the same country is called _____.
 (a) outsourcing
 (b) onshoring
 (c) nearshore outsourcing
 (d) onshore outsourcing
21. In which type of outsourcing payroll, are billing and HR outsourced?
 (a) Internal outsourcing
 (b) External outsourcing
 (c) Front-office outsourcing
 (d) Back-office outsourcing
22. In which type of outsourcing is customer service outsourced?
 (a) Internal outsourcing
 (b) External outsourcing
 (c) Front-office outsourcing
 (d) Back-office outsourcing
23. Which of these is not a valid reason for outsourcing?
 (a) Reduce costs
 (b) Efficiency of tasks
 (c) Both of these
 (d) None of these
24. Equity research is done in _____.
 (a) BPO
 (b) BDP
 (c) KPO
 (d) PPO
25. Identify the disadvantages of KPOs and BPOs.
 (a) Inferior quality of output
 (b) Increased complexities
 (c) Both (a) and (b)
 (d) None of these
26. Select the most preferred countries for voice-based KPOs.
 (a) India
 (b) Chile
 (c) Mexico
 (d) All of these
27. Select the most preferred countries (after India) for BPOs.
 (a) Sri Lanka
 (b) Bangladesh
 (c) Philippines
 (d) All of these
28. Information systems include _____.
 (a) information and software
 (b) hardware and data
 (c) people
 (d) all of these
29. Which type of information system will small organizations usually not have?
 (a) TPS
 (b) MIS
 (c) OAS
 (d) DSS
30. At which level of information system do senior managers work?
 (a) EIS
 (b) DSS
 (c) TPS
 (d) KMS
31. At which level of information system are operational decisions taken?
 (a) EIS
 (b) DSS
 (c) TPS
 (d) KMS
32. Which of the following is not a part of OAS?
 (a) EDI
 (b) Voice email
 (c) Hardware
 (d) Office building
33. Electronic data interchange is a type of _____.
 (a) hardware
 (b) software
 (c) technology
 (d) none of these
34. At which level of information system is collecting customers' feedback done?
 (a) EIS
 (b) DSS
 (c) TPS
 (d) KMS
35. At which level of information system are different types of reports generated?
 (a) MIS
 (b) DSS
 (c) TPS
 (d) KMS
36. In which type of report is aggregated data displayed?
 (a) Detailed
 (c) Exception

- ## Answers to Multiple-choice Questions

1. (a) 2. (b) 3. (c) 4. (c) 5. (c) 6. (a) 7. (a) 8. (b) 9. (d) 10. (a)
11. (d) 12. (d) 13. (a) 14. (c) 15. (c) 16. (c) 17. (c) 18. (a) 19. (a) 20. (d)
21. (d) 22. (c) 23. (d) 24. (c) 25. (c) 26. (d) 27. (d) 28. (d) 29. (c) 30. (b)
31. (c) 32. (d) 33. (c) 34. (c) 35. (a) 36. (b) 37. (c) 38. (a) 39. (b) 40. (d)
41. (d) 42. (d)

Data Organization and Database Management System

CHAPTER

2

Syllabus Mapping	Unit
(a) Data organisation: Character, field, record, file and database, types of data processing systems [serial, batch, real-time, online, centralized, distributed], file organizations [sequential, direct, indexed-sequential, relative], traditional file organisation vs. database file organisation. (b) Database management system: Concept of database management system (DBMS), definition, importance of DBMS, important terms of database [entity, attribute, keys, primary, foreign and candidate, referential integrity, table, views, data dictionary], types of database [hierarchical, network, and relational], basic ideas of data warehouse and data mining (definition, importance, advantages, and disadvantages), big data analysis—concept.	Module I Unit 2

2.1 INTRODUCTION

Today, most applications collect a huge amount of data. Across workplaces, a lot of data is collected in one form or the other. For example, when we take admission in a college, we provide data in the form of our name, address, phone number, course in which to seek admission, aggregate of marks obtained in the last examination, and so on. If we go to a bank to get an account opened, we are always required to provide a lot of data. All this information was traditionally stored on paper but handling these documents had never been easy.

Similarly, scientific experiments and satellites also generate enormous amounts of data. Therefore, in order to efficiently analyse the data collected from different sources, it has become a necessity to store the data in computers in the form of a file.

In computer terminology, a file is a block of useful data, which is available to a computer program and is usually stored on a persistent storage medium like a hard disk. This ensures the availability of the file for future use. Nowadays, files stored on computers are a good alternative to store documents that were once stored in offices and libraries.

2.2 DATA ORGANIZATION

Every file contains data. Data can be organized in a hierarchy to present a systematic organization. Data hierarchy includes terms such as fields, records, files, and database. These terms can be defined as follows:

- A data field is an elementary unit that stores a single fact. It is usually characterized by its type and size. For example, a student's name is a data field that stores the name of the student. This field belongs to the type called 'characters' and its size can be set to a maximum of 20 or 30 characters, depending on the need.

- A record is a collection of related data fields and is seen as a single unit from the application point of view. For example, a student's record may contain data fields such as the name of the student, address, phone number, roll number and marks obtained, and so on.
- A file is a collection of related records. For example, if there are 60 students in a class, then there will be 60 records of the students. All these related records are stored in a student's file. Similarly, we can have a file of all employees working in an organization, a file of all the customers of a company, a file of all the suppliers, and so on.
- A directory stores information of related files. A directory organizes information so that the user can find it easily. For example, consider Figure 2.1 that shows how multiple related files are stored in a student directory.

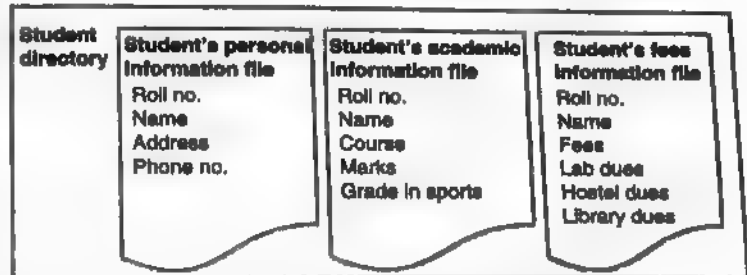


Figure 2.1 Directory, file, and attributes

2.3 DATA PROCESSING SYSTEM

The most important factor of a transaction system is its availability. Here, availability is defined as the fraction of time a system is up and running for processing. In simple terms, we can say that a system is available if it is not down because of hardware or software failures, operator errors, preventative maintenance, power failures, etc.

We can categorize data processing system with respect to the expected availability of the system. For example, a transaction processing system (TPS) installed at an airline reservation counter needs to be available 24×7 , but the TPS in a data entry operator who works from 9 am to 5 pm needs to be available only during the office hours. Therefore, based on how computers are being used for processing data and their availability required, we can have different types of data processing systems.

2.3.1 Batch Processing System

A batch is a set of requests that are processed together, often long after the requests were submitted. A batch can take minutes, hours, or even days to execute. Correspondingly, in batch processing systems, requests, jobs, or transactions to be performed by the computer are submitted to the computer and the results are obtained after some time. These types of systems were popular in the 1960s and the 1970s. Even today, they are used in TPS to execute the batches after office hours. A batch processing system would execute each batch as a sequence of transactions, one transaction at a time. For example, the value of a stock market portfolio can be calculated a day after the close of financial markets on all Saturdays, calculating telephone bill for customers on a monthly basis, and calculating tax annually.



The main advantage of batch processing systems is their throughput—the amount of work performed per unit of time. They are applicable where quick response time is not desirable.

Batch processing is usually used in systems that need to process a large amount of data in a specific way without user intervention in an automated manner. Data is collected during the working hours

and then the batch processed at the end of the working day when computing resources are not needed for other tasks. Generally, batch processing is used to perform repetitive tasks on large volumes of data.

2.3.2 Online Processing

Online data processing is used in systems that need to provide interactive computation. The moment data enters the system or a transaction needs to be processed, the system performs the processing without any delay. Unlike batch processing, the system does not wait to collect some additional data over a period of time. For example, in a retail showroom, we usually go and buy a product that has a barcode. The barcode reader reads the barcode, gives information about the product such as name, price, and so on, and deducts the product from available inventory. However, in a batch processing system, all inventory records are collected over a period of time—say over a week, and then the available inventory is updated after a week.

Some other examples of online processing include airline, bus, or railway reservation in which travellers would like to know whether a seat is immediately available or not. In case they make the booking, the number of available seats must immediately get deducted. Another example is that of ATM machines, which allow account holders to withdraw the required money. Game playing is also a field that uses online processing.

As these processes require users to supply input and system to immediately process it, online processing is also called interactive processing.



Online processing is faster and is used for processing continuous data.

2.3.3 Real-time Processing

In real-time processing, which is a subset of online processing, users submit requests to perform some transaction(s) that must be completed before the specified deadline (which is usually very near). Generally, real-time processing systems are used for collecting data from a satellite or from an instrument that continuously generates data. If they lose some input, they ignore the loss and keep on running. Users working with real-time systems expect a response in one or two seconds. The major advantage of real-time processing is that it automatically takes data from sensors and processes it without user intervention. Moreover, whenever a rapid reaction is required due to some sort of change, it immediately takes the necessary action.

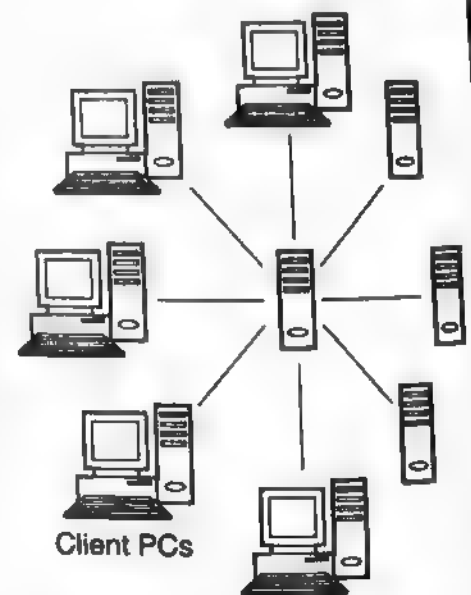


Figure 2.2 Centralized processing

2.3.4 Serial Processing

In serial processing, transactions are performed sequentially. This means that there is an explicit order in which transactions occur. It is applied in systems where one transaction must be completed before the next one begins. For example, before calculating the rank-wise list of candidates who appeared in an entrance exam, the individual result of each candidate must have been processed.

2.3.5 Centralized Processing

A central computer is a large computer that manages the resources of an organization and shares it with other computers in the system. Correspondingly, in centralized data processing, all data processing operations and calculations are performed by the central computer and different terminals or computers are connected to the central computer for sending requests and receiving outputs as shown in Figure 2.2.

Besides performing data processing operations, the central computer also manages the central data storage unit that stores the data needed by all the applications. The system administrator plays a vital role in centralized processing as he is responsible for authorizing user access, supporting users operations, and security of the system, thereby providing centralized control. Centralized systems are easy to manage and require less personnel costs.

For example, an airline ticket reservation service is a centralized system in which a central computer serves computers on several reservation counters. The central computer stores data about all flights, schedules, and vacancies. While it processes all reservation requests, the computer terminals at the counters are responsible for sending requests and printing the ticket.



All applications and data are mounted in the central computer, which is managed by a system administrator to provide a centralized control.

2.3.6 Decentralized or Distributed Processing

Nowadays, many organizations are moving from centralized processing to decentralized data processing (DDP). In this type of processing, relatively smaller computers located at different places in an organization are connected to each other with or without a central authority. The computers at each location are responsible for satisfying the specific needs of local users.

A popular phrase says, 'Do not put all your eggs in the same basket'. DDP implements this concept and therefore, distributes data and applications among several small computers. This also ensures that any failure of hardware, software, or personnel has minimum effect on the overall functioning of the system.

When a user requests for a service, the local computer is approached first. If the service can be provided by it, users can get a quick response; otherwise, the request is transferred to another computer. The major advantages of DDP include efficient utilization of resources and effective implementation of security and privacy measures. An organization need not invest in extra infrastructure in one go. Rather, it has the flexibility for gradual growth in hardware and software.

However, the downside of DDP is that it lacks standardization of data structures, resulting in problems of incompatibility and duplication of data. Since data is available on different computers, it is difficult to update data and exercise control over it.

A typical example of DDP is its use in banks where applications and data are split across different branches. Each branch processes data regarding its transactions and passes the summary information to the zonal office.

2.4 FILE ORGANIZATION

The particular organization most suitable for any application usually depends on factors such as the kind of external storage available, types of queries allowed, number of keys, mode of retrieval, and the mode of update.

We know that a file is a collection of related records. The main issue in file management is the way in which the records are organized inside the file because organization of records heavily affects

system performances. Here, by using the term 'organization of records', we mean the logical arrangement of the records in the file. For example, the organization is based on their ordering or the placement of related records close to each other in the file and not the physical layout of the file as stored on a storage media.

Choosing an appropriate file organization is a design decision; hence, it must be done keeping the priority of achieving good performance with respect to the most likely usage of the file. Therefore, the following considerations should be kept in mind before selecting an appropriate file organization method:

- Rapid access to one or more records
- Ease of inserting, updating, or deleting one or more records without disrupting the speed of accessing record(s)
- Efficient storage of records
- Using redundancy to ensure data integrity

Although one may find that these requirements are in contrast with each other, it is the designer's job to find a good compromise among them and get an adequate solution to the problem at hand. For example, ease of addition of records can be compromised to get fast access of data.

The four fundamental file organization techniques are as follows:

- Sequential
- Relative
- Indexed sequential
- Multi-key

These file organization techniques basically differ from each other in the way they perform the physical ordering of the records in the storage and the set of operations necessary to find particular records.

However, the most appropriate file organization depends on the operational characteristics of the storage medium. While some devices allow direct access to a particular record without accessing prior records stored on the device, others allow only sequential access to records. For example, while magnetic disks are direct access storage devices, magnetic tapes are sequential storage devices.

2.4.1 Sequential Organization

Sequential file organization is the most basic way to organize records. A sequentially organized file stores records in the order in which they are entered. The first record that is entered is written as the first record in the file, the second record entered is written as the second record in the file, and so on. As a result, new records are added only at the end of the file.

Sequential files can be read only sequentially, starting with the first record in the file. Sequential file organization is the most basic way to organize a large collection of records in a file. Figure 2.3 shows n records numbered from 0 to $n - 1$ are stored in a sequential file.

Once we store records in the file, we cannot make these records shorter or longer. We cannot even delete the records from a sequential file. However, a file can be updated only if the length does not change. In case, we need to delete or update one or more records, we have to replace the records by creating a new file.

While updating the sequentially stored data using a new file, the records are copied to the point where amendment is required. The changes are then made and copied into the new file. Following this, the remaining records in the original file are copied to the new file. Addition of new records in the sequential file is also done in a similar manner. When new records are added, all records till an appropriate point are shifted to make space for the new record. Inversely, when a record is deleted, the file space must be compressed by the shifting of records.

Hence, we see that updating records necessitates the creation of a new file thereby making it an expensive operation. However, to reduce the cost per update, all updates to data stored in a sequential file are batched, sorted in the order of the sequential file, and then used to update the file in a single pass. Such a file that contains the updates that have to be made to the sequential data file is known as the transaction file.

When updating a sequential data file using a transaction file, records requiring no changes are copied directly to a new file or the new master file; records requiring one or more changes are written into the new master file only after all necessary changes have been made. New records are inserted in the proper sequence in the new file and records to be deleted are simply not copied to the new master file.

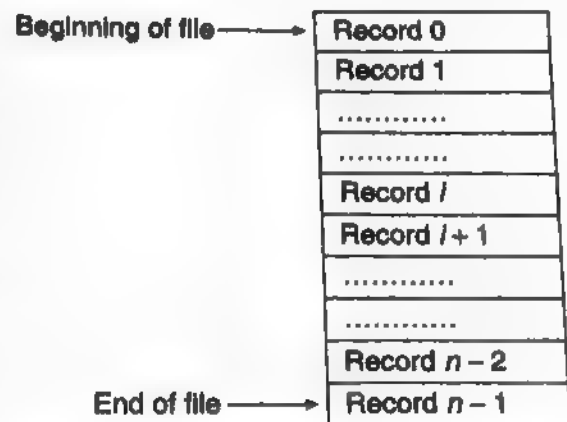


Figure 2.3 Sequential file organization

In sequential file organization, all records have the same size and the same field format, and every field has a fixed size. The records in the file are sorted based on the value of one field or a combination of two or more fields. This field is known as the key. Every key uniquely identifies the records in the file. Therefore, every record has a different value for the key field. Records can be sorted in ascending or descending order.

Sequential files are usually used for report printing or sequential reads of large amount of data which some programs prefer to do. An example is the payroll processing of all the employees of the organization. Sequential files can be easily stored both on disk and tape.

The processing of a sequential file is conceptually simple but inefficient for random access. However, if an application demands strictly sequential access to the data, then a sequential file is suitable. A sequential file could be stored on a sequential storage device like a magnetic tape.

Table 2.1 summarizes the features, advantages, and disadvantages of sequential file organization.

Table 2.1 Features, merits, and demerits of sequential file organization

Features	Advantages	Disadvantages
<ul style="list-style-type: none"> Records are written in the order in which they were entered. Records are read and written sequentially. Deleting or updating one or more records calls for replacing the original file with a new file that contains the desired changes. Records have the same size and the same field format. Records are sorted on a key value. It is usually used for report printing or sequential reads. 	<ul style="list-style-type: none"> It is simple and easy to handle. No extra overheads are involved. Sequential files can be stored on magnetic disks as well as magnetic tapes. It works well for batch-oriented applications. 	<ul style="list-style-type: none"> Records can be read only sequentially. If the ith record has to be read, then all $i - 1$ records must be read. It does not support for updating records in the same file. A new file has to be created and the original file has to be replaced with the new file that contains desired changes. It cannot be used for interactive applications.

2.4.2 Relative File Organization

Relative file organization provides an effective way to access individual records directly. In relative file organization, records in a file are ordered by their relative key. This means that the record number represents the location of the record relative to the beginning of the file. The record numbers start from 0 to $n - 1$, where n is the number of records in the file. For example, the record with record number 0 is basically the first record in the file. The records in a relative file are of fixed length.

Therefore, in relative files, records are organized on ascending relative record number. A relative file can better be thought of as a single dimension table stored on disk, in which the relative record number is the index of the table.

Relative files can be used for both random accesses of data as well as for sequential access. For sequential access, records are simply read one after another.

Relative files provide support for only one key, that is, the relative record number. This key must be numeric and must take a value between 0 and the current highest relative record number -1 . This means enough space must be allocated for the file to contain records with relative record numbers between 0 and the highest record number -1 . For example, if the highest relative record number is 1,000, then space must be allocated to store 1,000 records in the file.

Figure 2.4 shows a schematic representation of a relative file which has been allocated enough space to store 100 records. Although it has space to accommodate 100 records, not all the locations are occupied. The locations marked free are yet to store records in them. Therefore, every location in the table either stores a record or is marked as FREE.

Relative file organization provides random access of data by directly jumping to the record which has to be accessed. For example, if the records are of fixed length and we know that each record occupies 20 bytes and the base address of the file is 1,000, then, any record i can be accessed using the following formula.

$$\text{Address of } i\text{th record} = \text{base_address} + (i - 1) * \text{record_length}$$

Therefore, if we have to access the fifth record, then the address of the fifth record can be given as

$$\begin{aligned} & 1000 + (5 - 1) \times 20 \\ & = 1000 + 80 \\ & = 1080 \end{aligned}$$

Note that the base address of the file means, starting address of the file. We took $i - 1$ in the formula because, record numbers start from 0 rather than 1. Therefore, the fifth record is actually the fourth record. Table 2.2 summarizes the features, advantages, and disadvantages of relative file organization.

Table 2.2 Relative file organization

Features	Advantages	Disadvantages
<ul style="list-style-type: none"> It provides an effective way to access individual records. The record number represents the location of the record relative to the beginning of the file. 	<ul style="list-style-type: none"> Processing is easy. If the relative record number of the record that has to be accessed is known, then the record can be accessed instantaneously. 	<ul style="list-style-type: none"> Use of relative files is restricted to disk devices.

Relative record number	Records stored in memory
0	Record 0
1	Record 1
2	FREE
3	FREE
4	Record 4
.....
98	FREE
99	Record 99

Figure 2.4 Relative file organization

Features	Advantages	Disadvantages
<ul style="list-style-type: none"> The records in a relative file are of fixed length. Relative files can be used for both random accesses of data as well as for sequential access. Every location in the table either stores a record or is marked as FREE. 	<ul style="list-style-type: none"> Random access of records makes access to relative files fast. It allows deletions and updates in the same file. It provides random as well as sequential access of records with low overhead. New records can be easily added in the free locations based on the relative record number of the record to be inserted, It is well-suited for interactive applications. 	<ul style="list-style-type: none"> Records can be of only a fixed length. For random access of records, the relative record number must be known in advance.

2.4.3 Indexed Sequential File Organization

Indexed sequential file organization stores data for fast retrieval. The records in an indexed sequential file are of fixed length. In an indexed sequential file, every record is uniquely identified by a key field. We maintain a table known as index table or an index file that stores record number and the address of the record in the file. For every file, we have an index table. This type of file organization is called an indexed sequential file organization because the records may be stored anywhere physically but the index table stores address of those records.

The i th entry in the index table points to the i th record of the file. Initially, when the file is created, each entry in the index table contains NULL. When the i th record of the file is written, any free space is obtained from the free space manager and its address is stored in the i th location of the index table. Look at the Figure 2.5 which shows this scheme.

If one has to read the fourth record, there is no need to access the first three records. The address of the fourth record can be obtained from the index table and the record can be straightaway read from the specified address (742, in our example). Conceptually, the index sequential file organization can be visualized as shown in Figure 2.6.

The indexed sequential file uses the concept of sequential files as well as that of relative files. While the index table is read sequentially to find the address of the desired record, a direct access is made to the address of the specified record in order to access it randomly.

Record number	Address of the record
1	765
2	27
3	876
4	742
5	NULL
6	NULL
7	NULL
8	NULL
9	NULL

Figure 2.5 Indexed sequential file organization

Record number	Pointer to the record	
1		Record
2		Record
3		Record
4		Record
5	NULL	
6	NULL	
7	NULL	
8	NULL	
9	NULL	

Figure 2.6 Indexed sequential file organization

Indexed sequential files perform well in situations where sequential access as well as random access is made to the data. They can be stored only on devices that support random access of data, for example, magnetic disk.

Example Suppose a college stores the details of the students as an indexed sequential file. The file can be accessed in any of the following ways:

- Sequentially to print the aggregate marks obtained by each student in a particular course
- Randomly, for example, to modify the name of a particular student

Table 2.3 summarizes the features, advantages, and disadvantages of indexed sequential file organization.

Table 2.3 Indexed sequential file organization

Features	Advantages	Disadvantages
<ul style="list-style-type: none"> • It provides fast data retrieval. • The records are of fixed length. • Index table stores the address of the records in the file. • The ith entry in the index table points to the ith record of the file. • While the index table is read sequentially to find the address of the desired record, a direct access is made to the address of the specified record in order to access it randomly. • Indexed sequential files perform well in situations where sequential access as well as random access is made to the data. 	<ul style="list-style-type: none"> • The key improvement in is that the indexes are small and can be searched quickly, allowing the database to then access only the records it needs. • It supports applications that require both batch and interactive processing. • Records can be accessed sequentially as well as randomly. • It updates the records in the same file. 	<ul style="list-style-type: none"> • Indexed sequential files can be stored only on disks. • It needs extra space and overhead to store indices. • Handling these files is more complicated than handling sequential files. • It supports only fixed length records.

2.5 FILE-ORIENTED APPROACH

When computers were first used for business applications, a related group of records were stored in a file. In a file-oriented approach of data processing, each department has its own files that are specifically designed for those applications. Refer to Figure 2.7 in which there are three departments—academic, accounts, and library. Each of these departments maintains a student's personal details file in addition to the file which is necessary for their own application. Therefore, each application has a separate master file and its own set of personal files. The types of files maintained for each application include the following.

Transaction file It is a temporary file which stores data to be processed. Once the data is processed, it is permanently stored in the master file. More specifically, the transaction file is used to update the master file. For example, in a payroll application, the transaction file will be used to calculate the monthly salary of each employee, and for that it will contain data about number of hours worked, overtime done, and other details about each employee.

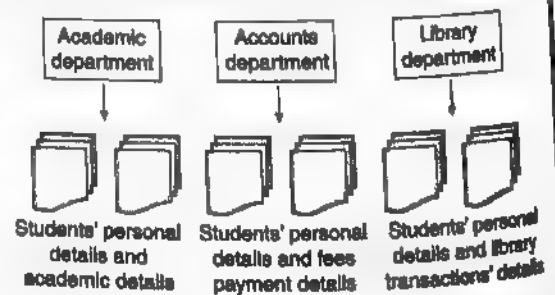


Figure 2.7 File-oriented approach

Master file
roll applica
scale, and a
it in the ma

Output file
be require
which can

Backup f
because o

2.6 D

A databa
and main
means th
For exam
ments l
there is
particu
presen

Applica
progra

Applic
progr

Figur

Ap

Dat
hav

Master file A master file contains all data relevant data for an application. For example, in case of payroll application the master file permanently stores details of each employee such as name, address, pay scale, and so on. The transaction file calculates the current month's salary of each employee and updates it in the master file to maintain a permanent record.

Output file When working with multiple programs for data processing, the output of one program may be required as an input of another program. Therefore, the first program stores its output in an output file which can then be used by other programs.

Backup file It is a copy of the file which the system keeps as a safety measure against loss of data because of inadvertent deletion. The original file can be restored from the backup file.

2.6 DATABASE APPROACH

A database is a collection of the related data organized in a way that allows users to easily access, update, and maintain the data. It stores non-redundant data that can be shared by different application systems. This means that only a single copy of data exists in the entire system. Non-redundancy ensures data consistency. For example, consider a scenario in which complete details of a student exist in files of three different departments like one in the accounts department, second in the academic department, and third in the library. If there is a change in a student's address and he reports it only in the academic department, then only that particular department will have the right data. This correct data will conflict with other copies of data that are present in other departments. This is called inconsistent data which occurred due to redundancy of data or

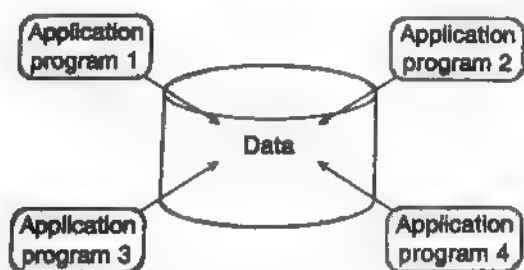


Figure 2.8 Database approach

existence of same data at multiple locations. Figure 2.8 shows the database approach in which all the programs access a single copy of data in the database.

Besides storing related data in an organized manner, a database also separates physical storage of data from application programs that access it. This is called data independence. In case of data independence, the programmer does not know the details of how the data is stored. These details are completely transparent to the users.

Applications of Database System

Database systems are widely used in almost all fields. Due to its numerous advantages, database systems have been successfully implemented for the following applications:

- They are used for ticket reservation of flights and trains.
- Airlines and railways use them for displaying the route and schedule of trains and flights.
- They are used in hospitals to store details of patients, their medical history, and other details.
- Financial institutions like banks use them for managing the records of customers, accounts, loans, and other transactions. Banks and other financial institutions also use databases to record purchases on credit cards for generating monthly bills.
- They are used in schools and colleges to store information about students.
- In organizations, databases are used to maintain records of all employees, their salaries, perks, taxes, and for generating salary cheques.
- In telecommunication departments, databases are used to store information about the telephone numbers, record of calls, for generating monthly bills, etc.

- Online shopping websites and other business activities use databases to store details of products, customers, transactions, and other payment details.
- The World Wide Web maintains enormous amount of textual and multimedia data in databases.
- They are used in online trading for storing information about sales and purchases of stocks and bonds.

Advantages of Database System

The advantages of using database to store data include the following:

- It reduction data redundancy because databases are shared rather than being used independently.
- It minimizes data inconsistencies.
- It maintains data integrity and quality.
- Information about the stored data including its meaning and interpretation can be saved in the data dictionary.
- The database management system (DBMS) takes care of fundamental operations that results in reduced cost of software development.
- The data stored is secured in a database as it includes security tools to control access.
- It is easy to search data by using keywords.
- Data can be shared among multiple branches of the same organization.

Data can be used to analyse trends such as which product is most popular among customers.

2.7 FILE-ORIENTED VS DATABASE-ORIENTED APPROACH

The file-oriented approach suffers from the following significant disadvantages.

Data redundancy In a file-oriented approach, if a particular data is required by two different applications, then it may be stored in two or more files. For example, in a college, the details of a student may be stored in the accounts department, academic department, and in the library.

Program or data dependency The major drawback of the file-oriented approach of data processing is that programs are dependent on the files and files depend on data. Therefore, when there was a change in the physical format of the file, the program also had to be changed. For example, if a new data field has to be added in the master file, then all programs that access the master file will have to be changed.

Lack of flexibility Since the data and programs are strongly coupled, most information retrieval requests would be predetermined. Users will be allowed to generate only certain types of reports and execute only a limited number of queries. Such systems are incapable of responding to un-anticipatory queries and some kind of investigative or trend analysis that has now become mandatory for any business to survive in the competitive world.

Difficulty in accessing data File-oriented data processing does not allow the needed data to be retrieved in a convenient and efficient manner. For example, if there is a need to make a list of all the customers who live within a particular postal code area and such a program is not available, then a new program will have to be created. If at a later point of time, there is a need to make a list of customers within a particular postal code earning above ₹10,00,000 per month, then either the existing program will have to be modified or a new program will have to be written. Therefore, with each new request for information either existing programs are modified or are written from scratch.

Data isolation Since data is scattered in different files in distinct departments and stored in different formats, writing new application programs to retrieve the appropriate data is difficult.

Data integrity issues Data stored in files must satisfy certain types of consistency constraints. For example, the student cannot issue more than three books. Developers enforce these constraints by adding appropriate code in the application programs. However, when new constraints have to be added, it is difficult to change the programs to enforce them. The problem is even more complex when constraints involve several data items from different files.

Data atomicity issues In case of a failure, data must be restored to the correct state that existed prior to the failure. It is difficult to accomplish this in the file oriented system. For example, a user runs a program to transfer ₹10,000 from his account to another account. While the money was transferred, the system failed. The money was deducted from his account but not credited to the other account resulting in an inconsistent database state. Therefore, the approach must either do it in entirety or not do it at all.

Security issues Not every user should be able to access all the data. For example, a librarian should not be allowed to access the salary of the teachers. Enforcing security constraints in a file-oriented system is difficult.

The disadvantages of using database to store data include the following.

Larger file size To support complex applications and functionalities, the DBMS is a big program that occupies a substantial amount of space on disk as well as in the internal memory.

Increased complexity A large number of complicated functions makes database processing a complex product. To efficiently work with databases, users must learn a lot to exploit it to its maximum potential.

Greater impact of failure In database approach, several users share the same data; so, if the system storing the data fails, then a number of users will be affected.

Difficulty in data recovery Recovering data in case of a catastrophe is not a trivial task. This is because several users may be trying to update a data simultaneously. If a catastrophe has occurred, the data is first restored to a prior state when it was last known to be correct. Then, any updates to the data since that time is redone. The more the number of users involved in updating the data, higher is the complexity of updating the database.

2.8 CONCEPT OF DATABASE MANAGEMENT SYSTEM

A database must store only successful transactions; a successful transaction has the property of ACID—atomic, consistent, isolated, and durable. These features can be explained as follows.

Atomicity When a transaction is performed, either all or none of its operations must be completed. This means a transaction that is successfully completed must commit (make changes to the database permanent) and the one that failed should abort (undo or erases database changes). For example, if A does an online money transfer to B, the transaction comprises two main operations—deducting amount from A's account and crediting B's account with that amount. Either both these operations should be performed or none.

Consistency Each transaction must preserve the consistency of the database. This means that after the transaction is complete, the database should have the correct information. For example, if there were ₹10,000 in A's account, then after paying ₹1,000 to B, the balance must be ₹9,000 and no other figure.

Isolation Each transaction must execute as if it were running alone. This is important to ensure that the result of running a set of transactions is the same as running one transaction at a time. This is called serializability and is implemented by locking the table. For example, if ₹1,000 has to be deducted from A's account and ₹3,000 has to be credited. Then, irrespective of whether the amount is deducted first or added, the end result should be that at the end of both the transactions, A's balance should be ₹12,000 (₹10,000 - ₹1,000 + ₹3,000).

Durability Durability means that the committed transactions are permanently stored on a storage device.

2.9 COMPONENTS OF DATABASE MANAGEMENT SYSTEM

An effective database system comprises four main components as follows: Refer to Figure 2.9

Data The most crucial component of a database is the data stored in it. Real-world applications demand storing and processing of large amount of data which are accessed by users through application programs. There are two types of data that a database stores—user's data and metadata. While user data contains data to supports user's applications, metadata, on the other hand, stores data about data. It describes the structure of the database and includes information regarding number of tables and table names, number of fields and field names, primary key fields, types of pre-defined queries, etc.

Hardware Hardware consists of the secondary storage devices on which the data is stored; input and output devices for receiving or giving data to users; and processor and main memory for processing the data in a fast and efficient manner. Since a database system can be either deployed on a desktop computer with a single user or on a mainframe computer to be shared among thousands of users, the hardware should be chosen to meet the database requirements.

Software Software consists of the DBMS which acts as a bridge between the user and the database. The DBMS software interacts with the user's application programs and database to insert, update, delete, and retrieve data. A DBMS is responsible for maintaining the integrity and security of stored data, and for recovering information in case of system failure. Some common examples of DBMS are MS ACCESS, SQL Server, Oracle, dBase, and FoxPro.

Users Users are the people who access the data from the database to perform their primary business responsibilities. The users may be clerical staff, managers, executives, etc. Based on the job profile, these users are either given full or partial access to the database data. Database users (refer to Figure 2.10) can be broadly classified into following categories:

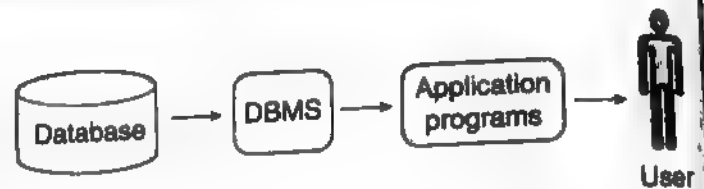


Figure 2.9 Components of a database-oriented approach

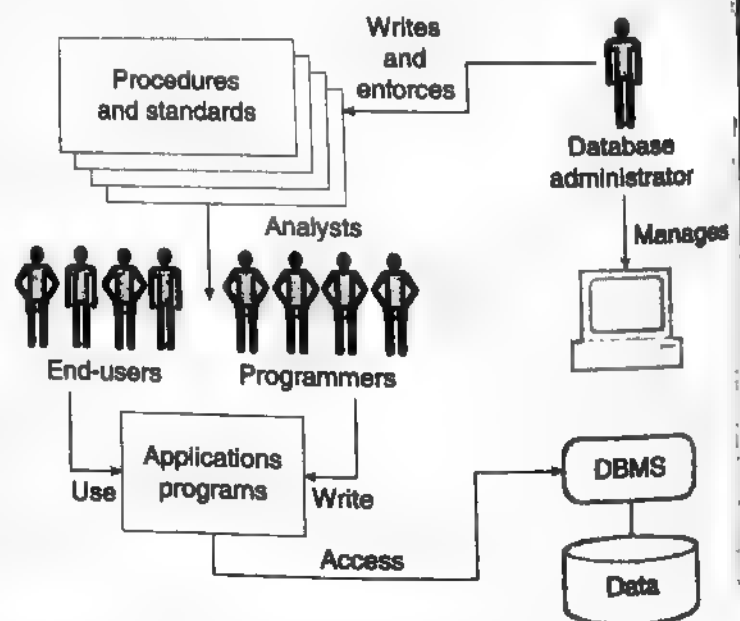


Figure 2.10 Database users

Application programmers An application programmer writes application programs to access, retrieve, update, delete, or add new data to the database. These programs are written using a high level language preferably, SQL which is a fourth generation language.

End-users End-users are those users who use the application programs developed by the application developers. They need not know about the database design, working, access mechanism, and other technical details of the database. Rather, they just use the underlying database to get their work done. End-users may be further categorized in two groups:

- Sophisticated end-users write their own queries to access and process the data stored in the database.
- Unsophisticated end-users interact with the system through an already written application program. They follow the instructions provided in the user interface to retrieve data from the database. For example, people working at railways reservation counters interact with the system through an already written application program. Another example of unsophisticated users is bank customers who use ATMs to draw money or check their balance.

Database administrator A database administrator (DBA) makes strategic and policy decisions regarding the data. He provides technical support for implementing these decisions and is responsible for overall control of the system at the technical level. He monitors and performs all activities related to database design, implementation, maintenance, and security. Besides administering the database, he also trains employees in database management and use.

System analyst System analysts identify end-users' requirements, plan solutions, and recommend hardware and software that best meet business goals. They play a major role in database design by writing end-users' requirements in a technical requirements definition document after conducting a technical and economical feasibility analysis of the identified requirements.

2.10 DATABASE VIEWS

A DBMS is a collection of interrelated data and a set of programs that enable multiple users to access, modify, share, and process it simultaneously. The main concern of a DBMS is to help users to efficiently retrieve data from the database. While ensuring efficiency, however, database developers often end up in designing complex data structure for data representation. Since a database is also used by unsophisticated end-users, this complexity must be hidden from them. Moreover, in the last section, we have seen that there are different categories of users with different needs for data. Therefore, a big responsibility of the DBMS is to provide users with an abstract view of the data to hide certain details of how the data is stored and maintained.

The DBA enforces data abstraction in database systems by defining data views at three levels—logical view, external view, and internal or physical view. These views provide a threefold benefit. First, it provides an abstract view of data. Second, it hides complexity from users. Third, it simplifies user interaction with the DBMS. These views can be discussed as follows.

Physical level This is the lowest level of abstractions that focus on how data is actually stored—the physical arrangement and location of data in the direct access storage devices (DASDs). It describes low-level complex data structures in detail. Database specialists use these details to design efficient algorithms to access data and make efficient use of storage and processing resources. The internal data representation is of no use to the end-users as they are just interested in using the information rather than in understanding the technical details of physical storage.

Logical level The logical view describes two things—first, what data is stored in database, and second, the relationships that exist among the data. It is the next higher level of abstraction that focuses on ease

of use. The users at this level, usually the DBAs, are least interested in knowing the physical-level complexity used to implement the structures.

External level Every user is concerned with a part of data in the database. External view is the next higher level of abstraction which is the sum total of user's views. Depending on the application, the DBMS supports multiple views of the external schema. The external level is the highest level of abstraction that hides all complexities to deal with the user's view of the database and exists only to simplify user interaction with system. Since all users and application programmers do not require the entire data stored in the database, the external level describes a part of the database for a particular group of users. It enables customization of data according to user's needs so that the same data can be seen by different users in different ways simultaneously. Besides customization of data, the external level also ensures security by hiding the parts of the database from certain users. The interrelationship among these three levels of abstraction is given in Figure 2.11.

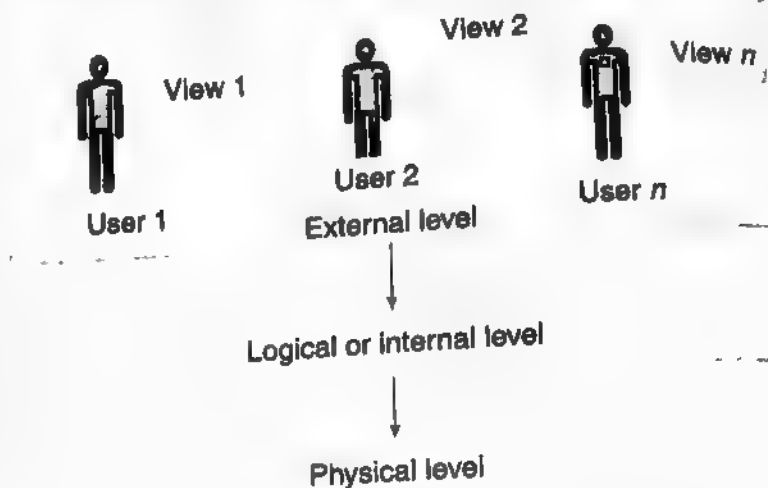


Figure 2.11 Interaction between different levels

2.11 THREE-SCHEMA ARCHITECTURE

A database schema is a layout of the database or a blueprint that outlines the manner in which the data is stored in the database. It describes the structure of a database in a formal language supported by the DBMS. For example, in a relational database (discussed in the next section), the schema defines the tables, its attributes or fields, and the relationships that exist between the fields and tables. Database schema is stored in a data dictionary or metadata in textual or in graphical format.



There may be multiple external schemas but there is only one logical and one physical schema.

Corresponding to the three levels of abstraction, there are three levels of database schema. This is called a three-level schema architecture or simply three-schema architecture. In the architecture, the three levels of schema are as follows.

Physical schema The physical schema describes the design of the database at the physical level. It incorporates the lowest level details that are hidden below the logical schema and describes the physical storage structure of the database. The physical schema can be easily implemented without affecting the application programs that are using the data. This schema is managed by the operating system under the direction of the DBMS.

Logical schema Logical schema describes the database design at the logical level. This schema is often used by the application developers and programmers to develop applications. It is considered to be the most important schema in terms of its effect on application programs.

External schema The external schema or the user's views are given at the external level. It describes a part of the database as per the user's requirements and hides the rest of the database from that user.

In three-schema architecture, each user group has its own external view. When a user or a user group requests to generate a new external view, the DBMS transforms the request specified at the external level into a request at the logical level, and then into a request at the physical level. When the user tries to access data from the database, the data must be first extracted and then presented in a format specified by the user. This process of transforming the requests and results between the three levels of schema architecture is called mapping.

The three-schema architecture ensures data independence, which is the main advantage of using this architecture. Data independence is the ability to modify the database schema at one level without affecting or without changing the schema at the other levels. It is of two types—logical data independence and physical data independence.

Logical data independence is the ability to change the logical schema of the database without affecting its external schemas or application programs. The logical schema may be changed due to any of the following reasons:

- Addition of new fields
- Addition of constraints
- Deletion of existing field(s)
- Removal of constraints

Physical data independence is the ability to change the internal schema without affecting the logical or external schema. An internal or the physical schema of the database may be changed due to the following reasons:

- Creating additional access structure
- Adding or deleting indexes
- Changing the storage structure



The physical schema is changed to improve the performance of the database system.

Generally, physical data independence is easier to achieve than logical data independence. This is because application programs depend on the logical structure of the database. Therefore, if there is any change in the logical structure of the database, some changes have to be made in the application programs too.

2.12 DATABASE MODELS

A model is a representation of reality that depicts real-world objects and their associations. A database model is, therefore, designed to provide the basic concepts and notations to enable application programmers, database designers, and end-users to communicate their understanding of the organizational data unambiguously and accurately. It describes three aspects—the data, the relationships that exist between data, and the constraints on that data. For this, the database model comprises the following components:

- A structural part that specifies rules for developing the database
- A manipulative part, which defines the types of operations that can be performed on the data, including those that change or access data or change the structure of the database
- Integrity rules to ensure that the data is accurate

The purpose of a database model is to represent data and to make it understandable. Some commonly used database models are as follows.

2.12.1 Hierarchical Model

The hierarchical data model organizes data in a tree structure. It forms a hierarchy of parent and child data segments. At the top is the root data element and below it there are the subordinate elements. Each of the subordinate elements may have its own subordinate elements, so on, and so forth. The tree can thus have multiple levels. The data elements in the hierarchical database have a parent-child relationship in which a parent can have n child data elements but a child data element has only one parent. This means there is a one-to- n mapping from the parent data element to the child data element. Figure 2.12 shows a hierarchical database of a college which has postgraduate as well as undergraduate students.

Hierarchical database allows related records to be grouped together to form a parent-child relationship. The parent data element collects information that is common to all the child records of the same group. This helps in reducing data redundancy in the database.

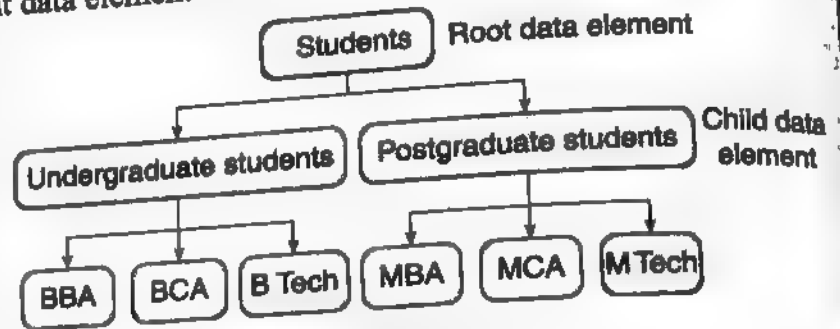


Figure 2.12 Hierarchical database

Advantages

The advantages of a hierarchical database are as follows:

- Many applications use a hierarchical database, for example, census data organized by state, within state by city, within city by zip code, etc.
- Its features lay the foundation for other data models.
- Since data access is quite predictable, data retrieval and update operations can be highly optimized by the DBMS.

Disadvantages

The disadvantages of the hierarchical database are as follows:

- It does not support flexible data access.
- Since data can be accessed only by following paths formed by branches of the tree structure, the database designer must determine all types of relationships among data elements and map them properly in a tree structure.
- It follows a complicated implementation pattern.
- It is difficult to manage.
- It lacks structural independence.
- It lacks standards.
- It is difficult to expand or modify the database. Any change in the structure requires substantial redesigning efforts.



The hierarchical data model consists of a set of records connected to one another through links. The link is an association between two or more records.

2.12.2 Network Model

The network model is a powerful yet complicated extension of the hierarchical database model. It has the following features:

- Data is represented by a collection of records.
- Relationships among data are represented by links.
- Each record of a particular record type represents a node.

However, in the network model, the nodes are not linked to each other in any hierarchy. While the hierarchical model organizes data in a tree structure, the network model, on the other hand, organizes it in the form of graphs. This means that, in the network model, parent data element can have many child data elements and a child data element can have many parent data elements. Therefore, the network model supports many-to-many relationships in data.

Figure 2.13 shows a network database. In the figure, subjects is the root data element. It has child data elements such as data warehousing, data structures, microprocessors, programming, and network. The child data elements have further child elements which are the courses in which the subjects are taught. As per the figure, data structures is taught in BSc, B Tech, and BCA. The figure exhibits many-to-many relationships among data elements. Note that M Tech has no parent data element. This means that none of the given subjects are taught in M Tech.

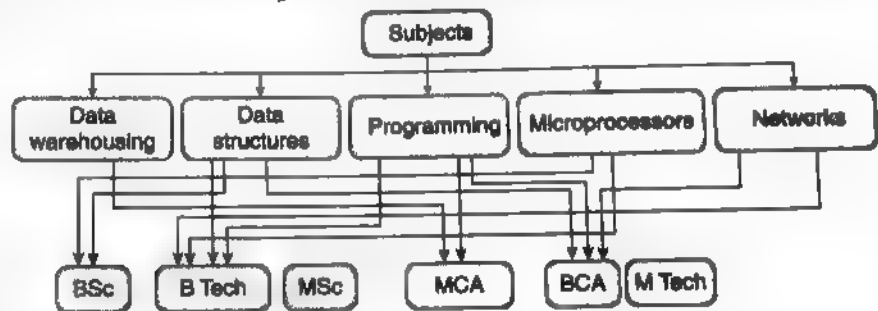


Figure 2.13 Network database



In the network model, a child can have multiple parents.

Advantages

The advantages of network model are as follows:

- It enables efficient and high-speed data retrieval.
- It enables the database designer to incorporate more relationship types.
- It provides easy data access.

Disadvantages

The disadvantages of network model are as follows:

- The database created using the network model has a complex structure.
- Database designers, administrators, and programmers must have a good knowledge of the internal data structures to access the data. Therefore, the model is not very user friendly.
- It is difficult to modify the database structure because any change will require the application programs to be modified before they can access data.

- It is difficult to maintain all the links; even a single broken link can be a big source of problems in the database.
- Since there is no strict bound on the number of relationships that can exist, the database design can become even more complex.

2.12.3 Relational Model

The relational data model was developed by E.F. Codd in 1970. It had no physical links among data elements as in the case of hierarchical and network models. In the relational model, data and the relationship that existed among those data are represented using tables, also known as relations. Therefore, a table is a collection of data records where each record contains the same fields. The main features of relational data model are as follows:

- A table consists of rows and columns.
- Each row stores a record.
- A record represents an entity.
- Each row is unique.
- Each column, also called field, represents an attribute of that entity.
- Each field has a unique name.
- A common attribute (and not physical links) is used to maintain a relationship between two tables.
- The sequence of rows and columns is insignificant.

Figure 2.14 shows the relational model consisting of two tables. The first table stores the roll number, name, address, and phone number of the students. The second table stores the roll number, course, and year of admission. Note that these two tables have a common field—RNo. Now if there is a query like 'In which course does Farhan study?' then the first table will be searched to obtain the roll number of Farhan. The roll number will then be used to search for a particular record in the second table. This means we are joining two tables based on a common attribute. The result of this query would return—BCA.

R no	Name	Phone no	Address
1	Anand	7838098765	Bank street
2	Bhaskar	9891123456	Park avenue
3	Dinkar	9818086420	Jai singh road
4	Farhan	9910035791	Copernicus marg
5	Harman	9999980654	Mall road

R no	Course	Year
1	BCA	2011
2	BSc	2012
3	BCA	2012
4	BCA	2013
5	BSc	2013

Figure 2.14 Relational model

Advantages

The advantages of relational model are as follows:

- Data model is simple to understand and use.
- Data access is simpler than other models.
- It is more programmer friendly and therefore, popular in both industrial and academic scenarios.
- It provides flexible data organization.

- Future enhancements to the database such as adding new table, row, field, and so on can be easily done.
- Relational model is close to the intuitive or logical model of real-life applications.

Disadvantages

Not all types of data can be represented using relational data model. Some examples of such data include multimedia, temporal, spatial, and unstructured data.

2.12.4 Object-oriented Data Model

The object-oriented paradigm can be applied to database technology to create an object-oriented data model. Such a model is a logical organization of the real-world objects or entities, constraints on them, and relationships among objects. The object-oriented database overcomes the shortcoming of relational, hierarchical, and network data models. This model can even store unstructured data such as images, audio, and video that can be used in complex applications. Some applications where they can be used are computer-aided design (CAD), computer-aided engineering (CAE), computer-aided manufacturing (CAM), computer-aided software engineering (CASE), expert systems, and multimedia systems. An object-oriented data model supports the following object-oriented concepts.

Object A real-world entity is uniformly modelled as an object having a unique identity and used to select an object to retrieve.

Attributes and methods Every object has a set of values for the attributes and methods which operate on the values of the object.

Message Methods of an object can be accessed or invoked from outside the object by explicitly passing messages to it.

Class It is a set of all the objects which share the same attributes and methods. An object must belong to only one class.

Inheritance A new class can be created from an existing class in such a way that the inherited class inherits all the attributes and methods of the existing class and may have additional attributes and methods.

The main advantage of the object-oriented data model is that it supports the development of a database application using an object-oriented programming language. This has made database implementation more convenient and efficient.

2.13 KEY TERMS

To understand the building blocks of relational databases, let us first get familiar with certain key terms.

Table Database is a collection of related tables. These tables are also known as relations. Therefore, databases that support the concept of relations are called relational databases. Each table is laid out in a tabular format consisting of rows and columns. While each row represents an entity or an object, columns are fields of the record. These fields represent the attributes of an entity.



Rows are also known as tuples.

Entities When designing databases, we need to first identify the entity or the subject areas for which information needs to be maintained. These subject areas will become tables. Entities can be of two types—one that has a physical existence and the other that has a conceptual existence (Figure 2.15). While the former can be a particular person, car, house, or employee, the latter can be a company or a course.

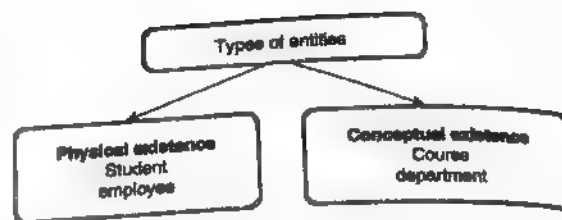


Figure 2.15 Types of entities

For example, in a college's database we can have a table Student to record his roll number, name, course, class teacher, marks, address, phone number, and other details.

Attributes Any entity is described using certain properties that are known as attributes. This means that an attribute is a property of the entity. For example, attributes of a student are name, roll number, course, marks, etc. Similarly, attributes of an employee are name, department, designation, manager, salary, etc.

While an entity is a uniquely identifiable object, an attribute, on the other hand, is an identified element within an entity.

Keys When designing tables, we need some mechanism to uniquely identify a particular record. For this purpose we use a key.

For example, in your class, each student has a unique roll number. Even if two students have the same name, they will have a different roll number. Similarly, when designing a table Student, the roll number can be used as a key. This key value will be used to search a record from a large number of records efficiently in less time. There are different types of keys as follows.

Candidate key A candidate key is a key that is eligible to become the primary key. For example, in a Student table with attributes such as roll_no, name, phone_no, course, marks, and address, one possible candidate key is the roll_no, and the second can be a combination of name and phone_no.

Therefore, we can say that a candidate key is an attribute or a combination of attributes that uniquely identifies an entity. Every table must have at least one candidate key although some tables can also have two or more.



A candidate key is a column or a combination of columns that satisfies all requirements of a primary key and therefore has the potential to become a primary key.

Primary key A primary key is an attribute, or combination of attributes, that allows each record to be identified uniquely. For example, if the Student table has attributes roll_no, university_enrolment_no, name, course, marks, etc., then either roll_no or university_enrolment_no can be made the primary key. Note that either of them but not both can be used as primary key as any one of them satisfies the purpose.

Alternate key If a table has more than one candidate key, then one of them is chosen as the primary key and all the other candidate

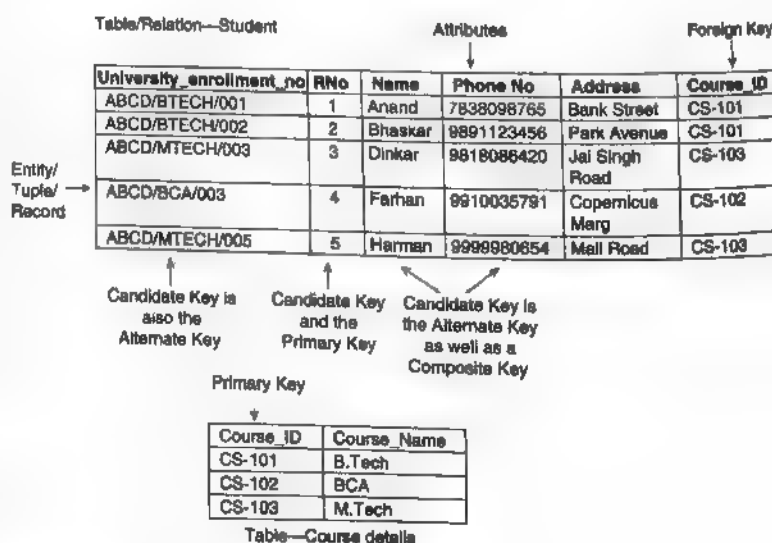


Figure 2.16 Tables and keys

keys are called alternate keys of that table. This means that primary key is also a candidate key and the remainder, if any, are called alternate keys.

For example, if we choose roll_no as the primary key, then university_enrolment_no and the combination of name + phone_no are the alternate keys as shown in Figure 2.16.

Composite key A composite key is a key made up of two or more attributes within a table to uniquely identify a record. For example, in the Student's table, the combination of name and phone_no forms a composite key.



It is always recommended to choose a simple key as the primary key rather than choosing a composite key.

Foreign key A foreign key is an attribute of a table that matches the primary key of another table. Such a key is specifically created to cross-reference tables. Foreign keys are used to restrict the domain of attribute(s) in the referencing relation and one or more records in the referencing relation can have the same value for the foreign key attribute.

For example, in a Student table, course_id is the foreign key. Course_id is the primary key of the table—Course Details. The Student's table can have limited values; that is only those specified in Course Details. Therefore, they are restricting the domain of attributes. Although there is a unique course_id in the Course Details table but in the Student table (referencing table), the course_id may have same values.



Foreign keys do not have a unique value in the referencing table.

From tables given in Figure 2.16, it is clear that primary keys are used to define the relationships among the tables. When primary key of one table is used as an attribute in another table, it becomes a foreign key in the other table. Moreover, a table can have more than one foreign key.

Referential integrity Foreign keys support the concept of referential integrity that ensures that relationships between tables remain consistent. When one table has a foreign key, the concept of referential integrity states the following.

- A record in a table with a foreign key cannot be added until there is a corresponding record in the linked table. This means that if there is no course MCA with course_id 4 in the Course Details table, then no student in the Student table can have course_id as 4.
- A record in a table with a foreign key is deleted if the foreign key, that is, the primary key of the other table is deleted. This means if CS-101 is deleted from the Course Details table, then all rows in the Student table with same course_id are also deleted. This property is called cascading delete.
- A record in a table with a foreign key is updated if the foreign key, that is, the primary key of the other table is changed. This means if CS-101 is updated to CS-111 in the Course Details table, it is still then all rows in Student table with same course_id are also updated. This property is called cascading update.

Data dictionary A data dictionary is a vital component of databases as it stores the metadata—data about data. Since it contains description of all the data stored in the database, it can be shared by several applications. Although a data dictionary is a crucial component, it is still invisible to the users and it

can be manipulated only by database administrators. Hiding data dictionary from users ensures that its contents are not destroyed, modified, or accessed by users with malicious intentions.

Data dictionary stores the structure of the data and usually stores the following information about data:

- Names of tables and their attributes
- Type of attributes in the table
- Length of each field in the table
- Domain of attributes—for example, date of birth should be between say 1914–2014
- Number of records in the table
- Primary key of the tables
- Foreign keys
- Access rights specifying who can use the database
- Source of data in the table
- Application programs that can use that data so that when the data is altered, the list of application programs that can be affected is generated easily

There are two types of data dictionaries—active data dictionary and passive data dictionary. While an active data dictionary is automatically updated as changes occur in the database, a passive data dictionary on the other hand must be manually updated.

Use of data dictionary The following are the uses of data dictionary:

- It is used by database designers, users, and administrators for getting useful information about data.
- It contains information about hardware and software.
- It maintains documentation on database design process.
- It contains information about data ownership.
- It stores information about relationships that exist among data.
- Database administrators use the data dictionary to control and maintain large databases.
- Database programmers use the data dictionary to develop programs and queries. They can also write procedures to manipulate data.
- It is used to maintain data integrity and accuracy.
- It allows users to know what data is available, in which format, and how it can be obtained.

2.14 RETRIEVING DATA THROUGH QUERIES

Database users can easily retrieve data from the database using the structured query language (SQL). The SELECT statement of the SQL allows users to query or retrieve data from a table in the database. The query may retrieve data from one or more columns in the table. The syntax of the SQL SELECT statement can be given as follows:

```
SELECT column_name(s) FROM table_name
```

where *table_name* is the name of the table from which the data has to be retrieved and *column_name* includes the name(s) of the columns from which the data has to be retrieved.

Consider the student table given in Figure 2.17.

RNo	Name	Phone No	Address
1	Anand	7838098765	Bank Street
2	Bhaskar	9891123456	Park Avenue
3	Dinkar	9818086420	Jai Singh Road
4	Farhan	9910035791	Copernicus Marg
5	Harman	9999980654	Mall Road

Figure 2.17 Student table

To select the names of all the students the following query can be executed on the database.

```
SELECT name FROM student;
```

The output of the query can be given as follows:

Anand
Bhaskar
Dinkar
Farhan
Harman

To select the names and phone number of all the students, the following query can be executed on the database.

```
SELECT name, phone-no FROM student;
```

The output of the query can be given as follows:

Anand	7838098765
Bhaskar	9891123456
Dinkar	9818086420
Farhan	9910035791
Harman	9999980654

The SELECT statement can be extended further to retrieve only specific and relevant data from the table leaving the rest. For this, it must be used with the WHERE clause. For example, if we want to see all the details of a particular student then using only the SELECT statement will retrieve the details of all the students in the class. This will increase the processing time for the query and the complexity to find the particular records amongst the record of the entire class. Therefore, the WHERE clause can be used to restrict the data that has to be retrieved. The syntax of the SELECT statement in conjunction with the WHERE clause can be given as follows:

The syntax for a WHERE clause with the SELECT statement is as follows:

```
SELECT column_name(s) FROM table_name  
WHERE condition;
```

Here, the condition is used to filter the rows retrieved from the table and gives only the ones those are actually required. Considering the table in Figure 2.17. The query to see the details of student with roll number 5 can be given as follows:

```
SELECT rno, name, phone-no, address FROM student  
WHERE rno = 5
```

2.15 DATA WAREHOUSE

Data warehouse is an information delivery system that integrates and transforms an organization's data into information to make it suitable for making strategic decisions. For building a data warehouse, current and historical data are pulled from the organization's databases and from other external sources and stored in the data warehouse.

W. H. Inmon, rightly called the father of data warehousing, defines data warehouse as 'a subject-oriented, integrated, non-volatile, [and] time-variant collection of data in support of management decisions'.

2.15.1 Subject-oriented Data

The transactional database focuses its attention on the day-to-day transactions that are part of the normal operation of the business. The data warehouse is concerned with the factors in the business environment that are driving those transactions.

In transactional systems, we store data by individual applications. In contrast, in the data warehouse, data is stored by subjects. These business subjects differ from enterprise to enterprise and are critical for the enterprise. For a manufacturing company, sales, shipments, and inventory are critical business subjects. For a retail store, sales at the check-out counter are a critical subject. In a data warehouse, there is no application flavour. The data in a data warehouse cut across applications. A particular subject may be involved in different types of transactions. For example, a customer appearing in the accounts receivable system may also be a supplier appearing in the accounts payable system. Each system has only part of the customer data. There is no single consolidated view of the whole organization.

Considering the way in which the decision-maker uses the data, this structure is very cumbersome. As decision-makers are interested in subjects, to get a complete picture of any one subject, they will have to access many tables within many applications. Data warehouse, therefore, gathers all of this data into one place so that the data for a particular subject is contained within one table.

2.15.2 Integrated Data

The warehouse, however, does more than gathering data. It derives its data from the transactional databases as shown in Figure 2.18. The integration process consists of two tasks—data cleansing and data transformation. The origin of the data is invisible to the users.

For proper decision-making, data in the data warehouse is taken from various applications in which data layout, character code representation, and field naming conventions could be different. For example, the account number in the Savings Account application could be eight bytes long, but only six bytes long in the Checking Account application.

Therefore, before putting this disparate data in the data warehouse, inconsistencies must be removed. Data elements must be standardized and the meanings of data names in each source application must be confirmed. This is called data transformation. Table 2.4 illustrates the importance of transformation during data integration. Moreover, the data must be cleansed to remove errors from the input data. If the cleansing process is faulty, the decision-maker does not trust the data and the warehouse fails.

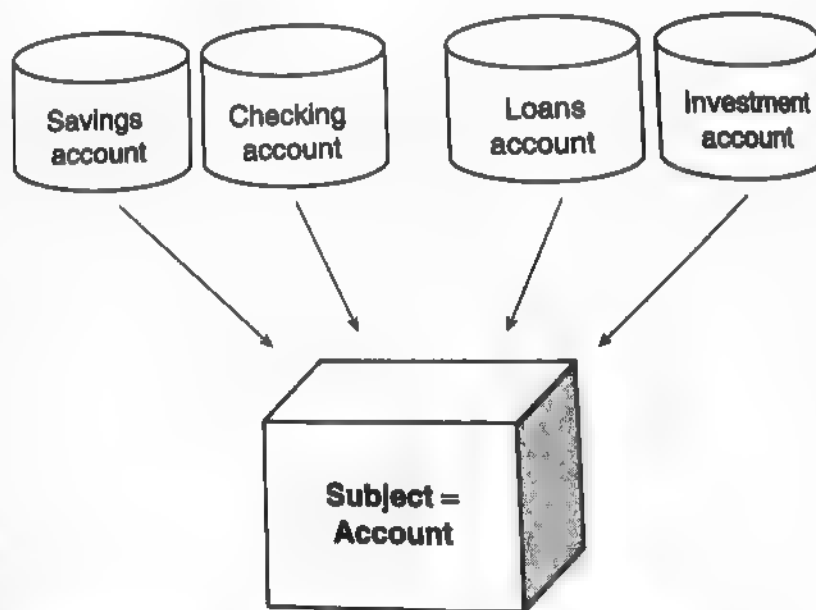


Figure 2.18 Data derived from transactional databases

Table 2

Descr
Enco
Unit
Co

2.1

Data
into
val
me
tw
on
2.
th
up
re
fr
h
d

2

R
c
s
s

Table 2.4 Integration Issues

	Sales voucher	Purchase order	Inventory
Description	Customer Name	CName	Customer Name
	IBM	IBM	International Business Machines
Encoding	Sex	Sex	Sex
	1 = Male 2 = Female	X = Male Y = Female	M = Male F = Female
Units	Cable length	Cable length	Cable length
	Centimetres	Metres	Inches
Coding	Key	Key	Key
	Character (10)	Integer	—

2.15.3 Non-volatile Data

Data from different systems are moved into the data warehouse at specific intervals. Depending on business requirements, these data movements take place twice a day, once a day, once a week, or once in two weeks. As shown in Figure 2.19, business transactions do not update the data in the data warehouse. They update only the transactional systems in real time. We add, change, or delete data

from a transactional system as each transaction happens but do not usually update the data in the warehouse. Data is not deleted in the data warehouse in real-time. This means that once data is stored the data warehouse, it is permanently stored there and is not altered. Hence, it is called non-volatile storage.

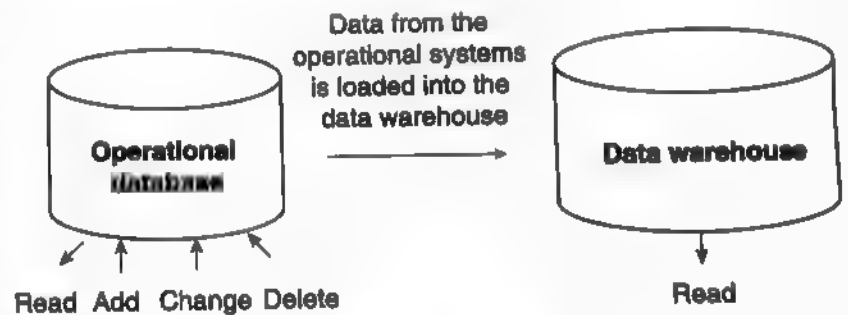


Figure 2.19 Non-volatility of data

2.15.4 Time-variant Data

For an operational system, the stored data contains the current values because these systems support day-to-day current operations. For example, in an order entry system, the status of an order is the current status of the order. On the other hand, the data in the data warehouse is meant for analysis and decision-making. If a user is looking at the buying pattern of a specific customer, the user needs data not only about the current purchase, but on past purchases as well.

A data warehouse, because of the very nature of its purpose, has to contain historical data and not just current values. Data is stored as snapshots over past and current periods. The time-variant nature of the data in a data warehouse (as shown in Figure 2.20) has the following features:

- Allows for analysis of the past

- Relates information to the present
- Enables forecasts for the future

Although a simple concept, it involves different functions—data extraction, loading the data, transforming the data, storing the data, and providing user interfaces, as shown in Figure 2.21.

- Time horizon—current from 60–90 days
- Key structure may/may not contain an element of time



- Time horizon—current as well as historical—5 to 10 years
- Key structure may/may not contain an element of time
- Sophisticated snapshots of data

Figure 2.20 Time-variant data

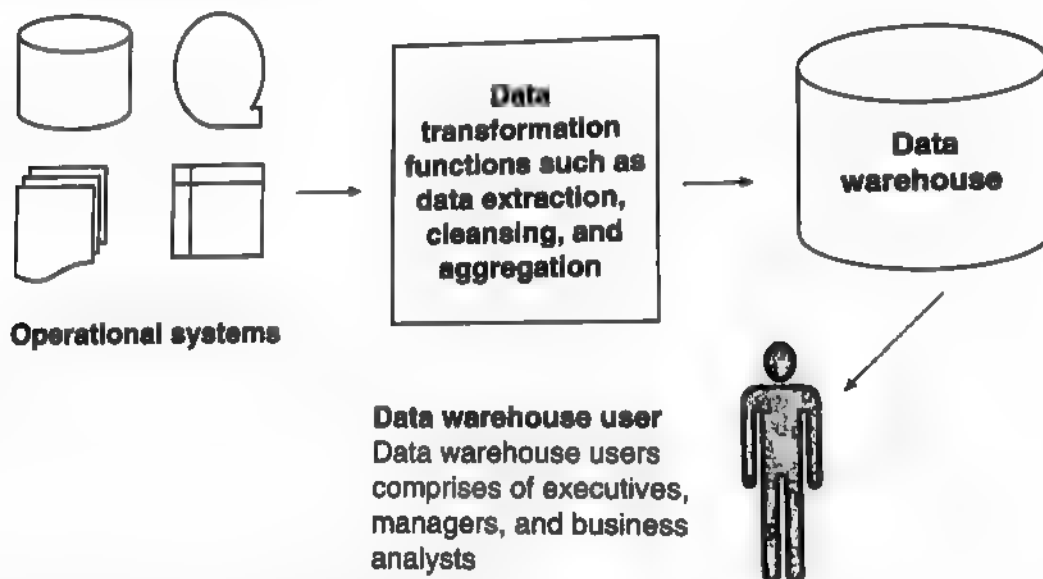


Figure 2.21 Data warehousing process

Table 2.5 shows the various uses of a data warehouse system.

Table 2.5 Uses of data warehouses

Industry	Uses of data warehousing
Retail	Customer loyalty and targeted marketing
Financial and banking	Risk management and fraud detection
Airlines	Route profitability and promotional schemes
Manufacturing	Cost reduction and resource management
Government	Manpower planning, development, and cost control

Table 2.6 summarizes the differences between the current operational system and the needed decision support system.

Table 2.6 Operational system vs decision support system

Attributes	Operational systems	Decision support/ informational system/data warehouse
Data content	Current values	Archived, summarized, derived
Data structure	Optimized for transactions	Optimized for complex queries
Access frequency	High	Medium to low
Access type	Read, update, delete	Read
Response time	Sub-seconds	Several seconds to minutes
User number	Large numbers	Relatively small number
Characteristics	Operational processing	Informational processing
Orientation	Transaction	Analysis
Users	Clerk, DBA, database professional	Knowledge worker, e.g., manager executives, analysts
Function	Day-to-day operations	Long-term informational requirements, decision support
Summarization	Highly detailed	Summarized
Unit of work	Short, simple transaction	Complex query
Number of records accessed	Tens	Millions
Database size	100 MB to GB	100 GB to TB

2.16 DATA MINING

Data mining refers to using a variety of techniques to identify nuggets of information or decision-making knowledge in the databases and extracting these in such a way that they can be put to use in areas such as decision support, prediction, forecasting, and estimation. The data is often voluminous, but it has low value and no direct use can be made of it. It is the hidden information that is useful.

As an analogy, imagine a very wide and very deep pit densely packed with some important material. You use a set of sophisticated drilling tools to dig and unravel the contents. You do not know what exactly you hope to get from your effort. Nothing may turn up, or you may be fortunate to find some real gold nuggets. You may discover this valuable treasure that you never knew was there in the first place. You

were not specifically looking for nuggets. You did not know they were there or if they ever existed. Look at Figure 2.22.

Now, as a change of scenario, replace the very wide and very deep pit with your data warehouse. Replace the material in the pit with the massive data content in your data warehouse and replace the drilling tools by data mining tools. The gold nuggets are the valuable pieces of information, such as patterns or relationships, you never knew existed in the data. In fact, you had applied the data mining tools to find something worthwhile you did not know existed. This is one aspect of data mining. Data mining is synonymous with knowledge discovery—discovering some aspect of knowledge you suspected to exist.

If knowledge discovery is one aspect of data mining, prediction is the other. Here, you are looking for a specific association with regard to an event or condition. For example, if you put a discount on color printers then some customers would like to buy the color printer as well while they are purchasing the computer.

Hence, data mining is a knowledge-discovery process. Data mining helps you understand the substance of the data in a special unsuspected way. It unearths patterns and trends in the raw data you never knew existed. Data mining centers on the automated discovery of new facts and relationships in data.

2.16.1 What can be Discovered?

The kinds of patterns that can be discovered depend upon the data mining tasks employed. By and large, there are two types of data mining tasks—descriptive data mining tasks that describe the general properties of the existing data, and predictive data mining tasks that attempt to do predictions based on inference on available data.

The data mining functionalities and the variety of knowledge they discover are briefly presented as follows.

Characterization Data characterization is a summarization of general features of objects in a target class and produces characteristic rules. The data relevant to a user-specified class are normally retrieved by a database query and run through a summarization module to extract the essence of the data at different levels of abstractions. For example, one may want to characterize the video library customers who regularly rent more than 30 movies a year.

Discrimination Data discrimination produces discriminant rules and is basically the comparison of the general features of objects between two classes referred to as the target class and the contrasting class. For example, one may want to compare the general characteristics of the customers who rented more than 30 movies in the last year with those whose rental account is lower than five. The techniques used for data discrimination are very similar to the techniques used for data characterization with the exception that data discrimination results include comparative measures.

Association analysis Association analysis is the discovery of association rules. It studies the frequency of items occurring together in transactional databases, and based on a threshold called support, identifies the frequent item sets. Another threshold, confidence, which is the conditional probability than an item

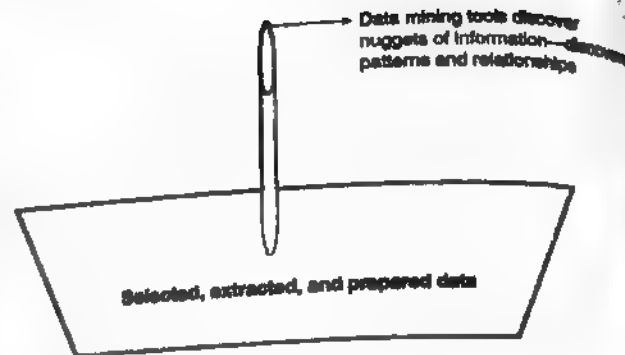


Figure 2.22 Data mining

appears
analysis
library
Classifi
classes
approac
The cla
sify ne
the cus
possibl
model
Predict
foreca
some
classifi
be for
howev
time
Clus
class
accep
not
max
simi
Out
are
be
don
Eve
dat
the
the
2.
Da
ple
op
th

appears in a transaction when another item appears, is used to pinpoint association rules. Association analysis is commonly used for market basket analysis. For example, it could be useful for the video library manager to know what movies are often rented together.

Classification Classification analysis or supervised classification is the organization of data in given classes. The classification uses given class labels to order the objects in the data collection. Classification approaches normally use a training set where all objects are already associated with known class labels. The classification algorithm learns from the training set and builds a model. The model is used to classify new objects. For example, after starting a credit policy, the video library managers could analyse the customers' behaviour vis-à-vis their credit, and label the customers who received credits with three possible labels—safe, risky, and very risky accordingly. The classification analysis would generate a model that could be used to either accept or reject credit requests in the future.

Prediction Prediction has attracted considerable attention given the potential implications of successful forecasting in a business context. There are two major types of predictions—one can either try to predict some unavailable data values or pending trends, or predict a class label for some data. The latter is tied to classification. Once a classification model is built based on a training set, the class label of an object can be foreseen based on the attribute values of the object and the attribute values of the classes. Prediction is, however, more often referred to the forecast of missing numerical values, or increase/ decrease trends in time related data. The major idea is to use a large number of past values to consider probable future values.

Clustering Similar to classification, clustering is the organization of data in classes. However, unlike classification, in clustering, class labels are unknown and it is up to the clustering algorithm to discover acceptable classes. Clustering is also called unsupervised classification, because the classification is not dictated by given class labels. There are many clustering approaches all based on the principle of maximizing the similarity between objects in a same class (intra-class similarity) and minimizing the similarity between objects of different classes (inter-class similarity).

Outlier analysis Outliers are data elements that cannot be grouped in a given class or cluster. They are also known as exceptions or surprises and are often very important to identify. While outliers can be considered noise and discarded in some applications, they can reveal important knowledge in other domains, and can thus be very significant and their analysis valuable.

Evolution and deviation analysis Evolution and deviation analysis pertain to the study of time related data that changes in time. Evolution analysis models evolutionary trends in data. Deviation analysis, on the other hand, considers differences between measured values and expected values, and attempts to find the cause of the deviations from the anticipated values.

2.17 DATA MINING AND DATA WAREHOUSE

Data mining fits well and plays a significant role in the data warehouse environment. A clean and complete data warehouse forms the bedrock for data mining and the data warehouse enables data mining operations to take place. The two technologies support each other. The following are some factors of this relationship:

- Data mining algorithms need large amounts of data, more so at the detailed level. Most data warehouses contain data at the lowest level of granularity.
- Data mining flourishes on integrated and cleansed data. If your ETL functions were carried out properly, your data warehouse contains such data, very suitable for data mining.

- The infrastructure for data warehouses is already robust, with parallel processing technology and powerful relational database systems. As such scalable hardware is already in place, no new investment is needed to support data mining.

2.17.1 Data Mining—A Data Warehouse Tool

Data mining is a technology that applies sophisticated and complex algorithms to analyse data and expose interesting information for analysis by decision-makers. Table 2.7 summarizes the differences between a typical data warehouse analysis tool and data mining tool.

Table 2.7 Differences between data warehouse analysis tool and data mining tool

Features	Data warehouse analysis tool	Data mining
Motivation for information request	What is happening in the enterprise?	Predict the future based on why this is happening
Data granularity	Summary data	Detailed transaction level data
Number of attributes	Small number of attributes	Many attributes
Sizes of datasets for the dimensions	Not large for each dimension	Usually very large for every dimension
Analysis approach	User driven and interactive	Data driven—automatic knowledge discovery

2.17.2 Applications of Data Mining

There is a wide variety of applications benefitting from data mining. The technology encompasses a rich collection of techniques that cover a wide range of commercial and non-commercial applications.

Non-commercial use of data mining is more pervasive in the research area. In oil exploration and research, data mining techniques discover locations suitable for drilling, based on the availability of potential mineral and oil deposits. Similarly, pattern discovery and matching techniques have military applications in providing assistance to identify targets. Medical research is a field where data mining helps researchers in examining correlations between diseases and patient's symptoms. Crime investigation agencies use data mining to connect criminal profiles to crimes. In astronomy and cosmology, data mining helps to predict cosmic events.

Apart from these areas, data mining technology has widespread applications in the commercial arena. The following is a list of applications of data mining across various areas in business.

Customer segmentation It is the most widely used application. Businesses use data mining to understand their customers. Cluster detection algorithms discover clusters of customers sharing the same information.

Market basket analysis This is a very useful application for retail. Association analysis algorithms uncover affinities between products that are bought together. Data mining also helps businesses to find customers to whom they can sell higher value items during an auction and notify them of an upcoming auction.

Risk management Insurance companies and mortgage businesses use data mining to uncover risks associated with potential customers.

Fraud detection Credit card companies use data mining to discover abnormal spending patterns of customers. Such patterns can expose fraudulent use of the cards.

Delinquency tracking Loan companies use the technology to track customers who are likely to default on repayments.

Demand prediction Retail and other businesses use data mining to match demand and supply trends to forecast demand for specific products.

2.17.3 Benefits of Data Mining

Data mining helps uncover the useful knowledge lying buried under mountains of data across companies. Otherwise, the data would have never been discovered and the benefits from using the discovered patterns and relationships would have never been realized.

In order to appreciate the enormous utility of data mining, let us enumerate the types of benefits of data mining actually realizable in real-world situations:

- In a manufacturing company, some employees of a department can be corrupted and there may be a variation between the purchase orders and the freight bills. Data mining detects criminal behaviour by uncovering patterns of orders and premature inventory reduction.
- A company can improve sales of a product by sending direct promotional discounts to its customers through a targeted campaign.
- A retail chain can improve its daily sales by rearranging the shelves based on discoveries of affinities of products that sell together.
- An airlines company can increase sales by discovering travelling patterns of frequent flyers.
- An insurance company can save large amount of money by detecting fraudulent claims.
- A manufacturer company can increase sales by forecasting sales of their products based on patterns discovered from the historical data stored in the data warehouse.
- A banking corporation with investment and financial services can prevent loss by detecting early warning signs for attrition in its checking account business.

2.17.4 Disadvantages of Data Mining

Though data mining technology helps in decision-making that is beneficial for business growth, one can come across several disadvantages of data mining systems. These disadvantages include the following:

Privacy issues With a rising number of people on social networking websites and increase in online shopping and other online activities, a large amount of data is being generated every day and stored for analysis. The companies are even collecting data whenever they have a new customer or when an existing customer buys a product. Just imagine, whenever a business is acquired by another company, all personal information it owns is then given to that company which may sell (or leak) this information. According to a news article, American Express had sold the details of the credit card purchases made by their customers.

Security issues Security is a big issue. Businesses store a lot of information about their employees and customers. This information includes personal details such as PAN, ADHAAR number, credit card number, birthday, salary, and address. All this data must be kept very safely. We often hear about hackers accessing and stealing the data of customers from big corporations such as Ford Motor Credit Company, Sony, and recently even from Indian banks. Even the US presidential polls in the year 2016 were said

to be manipulated. Hence, we need to understand that with so much personal and financial information available, identity theft is becoming a big problem.

Misuse of information/inaccurate information We are all afraid that information collected through data mining for ethical purposes can be misused. This information, if accessed by unethical people, can be used to discriminate against a group of people. Moreover, data mining techniques are not perfectly accurate. If these inaccurate results are used for decision-making, then the business in turn will have to face serious consequences.

Overwhelming amount of data For data mining systems, a large amount of data is collected. However, the main problem that arises during data collection is that the amount of data collected becomes overwhelming. The second problem is that some of the data may be irrelevant or even inaccurate. Therefore, it is very much essential to limit the data in some way and also to ensure that the collected data is cleaned (corrected) before being used by the data mining techniques.

Difficult to use Some of the data mining analytics software are difficult to operate and require the user to have knowledge-based training. Different data mining tools use different algorithms. Hence, the selection of the right data mining tool is very important to get appropriate results. To choose the right tool, a sound knowledge of the various available tools, features, and algorithms they employ is required. Usually, this knowledge is beyond the understanding of an average business analyst or knowledge worker. This is because data mining tools were especially designed for expert statisticians involved in data sciences.

This is a big challenge in the application of data mining techniques as an advanced level of analysis is reserved for the few, instead of for the masses. This prevents the full value of data mining from being exploited in the organization. Last but not the least, implementing data mining techniques in a company calls for investment of money for purchasing tools, hiring experts, and costs involved in maintenance, training and support, and others.

2.18 BIG DATA

Big data analytics is the process of collecting, organizing, and analysing large sets of data (called *big data*) to discover patterns and other useful information. It can help organizations to better understand the information contained within the data. Big data analysis allows users to evaluate large volumes of transaction data and other data sources that traditional systems would not be able to handle.

Big data is a term used for analysing data sets whose size or type is beyond the ability of traditional relational databases to capture, manage, and process with low-latency (less delay). The special features of data sets used in big data analysis include high volume (large amount), high velocity (frequently changing), or high variety (different types). This data comes from sensors, devices, video/audio, networks, log files, transactional applications, web, and social media. Much of the data is generated in real time and on a very large scale.

Analysing this big data allows data scientists or business users to make better and faster decisions using data that was previously inaccessible or unusable. Big data analytics uses advanced analytics techniques such as text mining, data mining, machine learning, predictive analytics, forecasting, data optimization, and natural language processing. Although these processes are separate, they are integrated for high-performance analytics.

Sophisticated software tools are used for big data analytics. Big data analysis outperforms data warehousing as it can easily handle large amount of unstructured and semi-structured data efficiently.

Furthermore, data warehouses may not be able to handle the processing demands posed by sets of big data that are updated frequently or even continually. This is especially important in the case of real-time data on stock trading or analysing the online activities of website visitors or the performance of mobile applications.

Big data analytics environments and technologies including Hadoop, MapReduce, and NoSQL databases can process huge data sets over clustered systems.



Big data analysis techniques can analyse large data sets of size ranging from few terabytes to zettabytes.

2.18.1 History and Evolution of Big Data Analytics

The history of big data analysis can be traced back to 1950s when people were not even aware of the term *big data*. In those days, people were analysing business numbers in a basic spreadsheet software to uncover insights and trends in the business.

Later, businesses switched to data warehousing and data mining techniques to gather information, run analytics, and unearth information that could be used for future decisions. However, today businesses have grown to an extent that they are seeking to identify insights for immediate decisions. The ability to work faster and stay agile gives organizations a competitive edge like never before.

2.18.2 Benefits of Big Data

Big data analytics helps organizations to harness their data and identify new opportunities. This leads to smarter business moves, more efficient operations, higher profits, and happier customers. Big data analysis helps business in the following ways:

Cost reduction Big data technologies such as Hadoop and cloud-based analytics give significant cost advantages when it comes to storing large amounts of data. This data can be analysed to identify more efficient ways of doing business.

Faster and better decision-making Because of the speed at which new data sources and types of data can be analysed, businesses are able to analyse information immediately and make decisions based on what they have learned.

New products and services The ability to map customer needs and satisfaction through analytics gives customers what they exactly want. Therefore, businesses are using big data analysis techniques to meet customer needs.



Big data analysts typically want the *knowledge* that comes from analysing the data.

2.18.3 Applications of Big Data

Researchers are using big data to do the following things:

- Decode human DNA in minutes
- Predict where terrorists plan to attack
- Determine which gene is most likely to cause certain diseases
- Determine the ads a user is most likely to respond to on Facebook

- Orange Mobile company in France used 2.5 billion records of its customers (keeping them anonymous) for analysis. These records included details on calls and text messages exchanged among 8 million users. Researchers analysed the data and proposed projects that showed two things. First, how to improve public safety by tracking cell phone data to map where people went after emergencies. Second, how to use cellular data to find out the source of a disease.

Summary

- When computers were first used for business applications, a related group of records were stored in a file.
- In batch processing systems, requests, jobs, or transactions to be performed by the computer are submitted to the computer and the results are obtained after some time.
- Online data processing is used in systems that need to provide interactive computations.
- In serial processing, transactions are performed sequentially.
- An effective database system comprises of four main components—data, hardware, software, and users.
- In real-time processing, which is a subset of online processing, users submit requests to perform some transaction(s) that must be completed before the specified deadline (which is usually very near).
- In centralized data processing, all data processing operations and calculations are performed by the central computer and different terminals (computers) are connected to the central computer for sending requests and receiving outputs.
- In decentralized data processing (DDP), relatively smaller computers located at different places in an organization are connected to each other with or without a central authority.
- A sequentially organized file stores records in the order in which they were entered.
- Relative file organization provides direct access to individual records.
- The main concern of database management system (DBMS) is to help users to efficiently retrieve data from the database.
- The database administrator enforces data abstraction in database systems by defining data views at three levels—logical view, external view, and internal or physical view.
- A database model describes three things—data, the relationships that exist between data, and the constraints on that data.
- The data definition module provides functions to define the structure of the data and the data manipulation module provides functions to perform operations such as inserting, searching, and deleting data in the database.
- Database users can easily retrieve data from the database using structured query language (SQL). The SELECT statement of SQL allows users to query or retrieve data from a table in the database.
- The query processor accepts user's queries and transforms them into a series of low level instructions. The report writer utilizes the output of query execution to display it in an easy to understand and interpret format.
- The hierarchical data model supports 1 to N mapping while the network model supports a many-to-many data mapping.
- A data dictionary is a vital component of databases as it stores the metadata—data about data.
- Data warehouse is an information delivery system that integrates and transforms the organization's data into information to make it suitable for making strategic decisions.
- Data mining is a technology that applies sophisticated and complex algorithms to analyse data and expose interesting information for analysis by decision-makers.
- Big Data analytics can help organizations to better understand the information contained within the data.

Glossary

Alternate key, then on and all the key of that

Batch A se often long

Big Data organizing Big Data information

Candidate primary k

Central resources computer

Compos of two o identify

Database a way t maintain

Database

DBMS program ify, sha

Direct

Field

Mu

1.

2.

Glossary

Alternate key If a table has more than one candidate key, then one of them is chosen as the primary key and all the other candidate keys are called alternate key of that table

Batch A set of requests that are processed together, often long after the requests were submitted

Big Data analytics The process of collecting, organizing, and analysing large sets of data (called Big Data) to discover patterns and other useful information

Candidate key A key that is eligible to become the primary key

Central computer A large computer that manages resources of an organization and shares it with other computers in the system

Composite key A composite key is a key made up of two or more attributes within a table to uniquely identify a record

Database A collection of related data organized in a way that allows users to easily access, update and maintain the data

Database schema A layout of the database

DBMS A collection of interrelated data and a set of programs that enables multiple users to access, modify, share, and process it simultaneously

Directory A collection of related files

Field An elementary unit that stores a single fact

File A collection of records

Foreign key An attribute of a table that matches the primary key of another table

Index table A table that stores record number and the address of the record in the file

Link An association between two or more records

Logical data independence The ability to change the logical schema of the database without affecting its external schemas or application programs

Mapping The process of transforming the requests and results between the three levels of schema architecture

Metadata Data about data

Physical data independence The ability to change the internal schema without affecting the logical or external schema

Query The statement written to retrieve information

Query language The part of the DML that retrieves the information from the database

Record A collection of related data fields that is seen as a single unit from the application point of view

Relational database A collection of related tables also known as relations. Therefore, databases that support the concept of relations are called relational databases

Table A collection of data records where each record contains the same fields

Multiple-choice Questions

1. Which of the following is an elementary unit that stores a single fact?

- | | |
|------------|---------------|
| (a) Field | (c) File |
| (b) Record | (d) Directory |

2. A collection of related data fields is known as a _____.

- | | |
|----------------|---------------|
| (a) data field | (c) file |
| (b) record | (d) directory |

3. A collection of related records is known as a _____.

- | | |
|-----------------|---------------|
| (a) file system | (c) file |
| (b) field | (d) directory |

4. A directory stores information of related _____.

- | | |
|-------------|------------------|
| (a) fields | (c) files |
| (b) records | (d) file systems |

5. The term that defines the fraction of time a system is up and running for processing is _____.
 (a) availability (c) accessibility
 (b) readability (d) usability
6. A _____ is a set of requests that are processed together, often long after the requests were submitted.
 (a) set (c) transaction
 (b) batch (d) record
7. Calculating the value of a stock market portfolio is an example of _____ processing.
 (a) batch (c) transactional
 (b) online (d) real-time
8. _____ processing is best in cases where a quick response time is not desirable.
 (a) Batch (c) Transactional
 (b) Online (d) Real-time
9. _____ processing is best in cases where interactive communication is desirable.
 (a) Batch (c) Transactional
 (b) Online (d) Real-time
10. Airline reservation system is an example of _____ processing.
 (a) batch (c) transactional
 (b) online (d) real-time
11. _____ processing is a subset of online processing.
 (a) Batch (c) Transactional
 (b) Serial (d) Real-time
12. Data is automatically taken from sensors and processed without user intervention in _____ processing.
 (a) batch (c) transactional
 (b) online (d) real-time
13. In _____ processing, transactions are performed sequentially.
 (a) transactional (c) serial
 (b) online (d) real-time
14. Efficient utilization of resources and effective implementation of security and privacy measures is difficult in _____ processing.
 (a) centralized (c) batch
 (b) decentralized (d) none of these
15. Processing in bank branches and airline reservation are respectively categorized as _____ processing.
 (a) centralized, decentralized
 (b) centralized, centralized
 (c) decentralized, decentralized
 (d) decentralized, centralized
16. Which of the following is not a file organization technique?
 (a) Indexed (c) Parallel
 (b) Sequential (d) Relative
17. _____ file organization stores records in the order in which they are entered.
 (a) Indexed (c) Multi-key
 (b) Sequential (d) Relative
18. Deleting and updating records in which kind of file requires replacing an old file with the new one?
 (a) Indexed (c) Multi-key
 (b) Sequential (d) Relative
19. All records having the same size and the same field format, and every field having a fixed size is a characteristic of which file organization?
 (a) Indexed (c) Multi-key
 (b) Sequential (d) Relative
20. In a sequential file, all records are sorted based on the _____ field.
 (a) name (c) key
 (b) ID (d) none of these
21. Payroll processing of all the employees of an organization can be done using a/an _____ file.
 (a) indexed (c) multi-key
 (b) sequential (d) relative
22. A/an _____ file does not support random access.
 (a) indexed (c) multi-key
 (b) sequential (d) relative
23. A/an _____ file does not support interactive applications.
 (a) indexed (c) multi-key
 (b) sequential (d) relative
24. A/an _____ random access.
 (a) indexed
 (b) relative
25. The record _____ the record in which type _____
 (a) Indexed
 (b) Sequential
26. A/an _____ application.
 (a) indexed
 (b) sequential
27. Records _____ file organization.
 (a) indexed
 (b) sequential
28. In the file _____ file is also _____
 (a) transactional
 (b) master
29. A/an _____ master file.
 (a) transactional
 (b) master
30. In the file _____ contain _____
 (a) transactional
 (b) master
31. In case _____ restore _____
 (a) transactional
 (b) master
32. A _____ data is _____ easily.
 (a) file
 (b) record
33. A data _____
 (a) operation
 (b) record
34. Separation _____ applica-

24. A/an _____ file can be used for both random access of data as well as for sequential access.
 (a) indexed (c) both of these
 (b) relative (d) none of these
25. The record number represents the location of the record relative to the beginning of the file in which type of file?
 (a) Indexed (c) Multi-key
 (b) Sequential (d) Relative
26. A/an _____ file is well-suited for interactive applications.
 (a) indexed (c) multi-key
 (b) sequential (d) relative
27. Records are of equal size in _____ type of file organization.
 (a) indexed (c) relative
 (b) sequential (d) all of these
28. In the file-oriented approach, the temporary file is also known as _____ file.
 (a) transaction (c) backup
 (b) master (d) output
29. A/an _____ file is used to update the master file.
 (a) transaction (c) backup
 (b) master (d) output
30. In the file-oriented approach, the _____ file contains all data relevant for an application.
 (a) transaction (c) backup
 (b) master (d) output
31. In case of data loss, the original file can be restored from the _____ file.
 (a) transaction (c) backup
 (b) master (d) output
32. A _____ is a collection of the related data organized in a way that allows users to easily access, update, and maintain the data.
 (a) field (c) file
 (b) record (d) database
33. A database stores _____ data.
 (a) organized (c) inconsistent
 (b) redundant (d) all of these
34. Separating physical storage of data from application programs that access it means _____.
 (a) organization
 (b) redundancy
 (c) inconsistency
 (d) data independence
35. The World Wide Web stores _____ data.
 (a) textual (c) both (a) and (b)
 (b) multi-media (d) none of these
36. Information about data in the database is stored in a _____.
 (a) file (c) data dictionary
 (b) directory (d) master file
37. Data in a database is searched for using _____.
 (a) words (c) fields
 (b) keywords (d) records
38. Data accessibility is the easiest and most efficient in a _____.
 (a) master file (c) output file
 (b) transaction file (d) database
39. Data stored in files must satisfy certain types of consistency constraints. This ensures data _____.
 (a) dependency (c) integrity
 (b) consistency (d) isolation
40. Data must be restored to the correct state that existed prior to the failure. This ensures _____.
 (a) atomicity (c) integrity
 (b) consistency (d) isolation
41. Not every user should be able to access all the data. This ensures _____.
 (a) atomicity (c) integrity
 (b) consistency (d) security
42. When a transaction is performed, either all or none of its operations must be completed. This ensures _____.
 (a) atomicity (c) integrity
 (b) consistency (d) security
43. After the transaction is complete, the database should have the correct information. This ensures _____.
 (a) atomicity (c) integrity
 (b) consistency (d) security

44. Committed transactions are permanently stored on a storage device. This ensures _____.
- (a) atomicity (c) integrity
(b) consistency (d) durability
45. The result of running a set of transactions is the same as running one transaction at a time. This feature is ensured by the property of data _____.
- (a) isolation (c) integrity
(b) consistency (d) durability
46. Which of the following is not a database software?
- (a) MS Access (c) MS Excel
(b) SQL Server (d) Oracle
47. Application programs to access, retrieve, update, delete, or add new data to the database is written using which language?
- (a) C (c) SQL
(b) C++ (d) Java
48. Technical details of databases may not be known by the _____.
- (a) end-users
(b) application programmers
(c) system analyst
(d) DBA
49. People working at railways reservation counters are an example of _____.
- (a) sophisticated end-users
(b) application programmers
(c) unsophisticated end-users
(d) DBA
50. Who makes the strategic and policy decisions regarding the data?
- (a) End-users
(b) Application programmers
(c) System analyst
(d) DBA
51. Who provides technical support for implementing these decisions and is responsible for overall control of the system at the technical level?
- (a) End-users
(b) Application programmers
(c) System analyst
(d) DBA
52. Training employees in database management and use is the work of the _____.
- (a) end-users
(b) application programmers
(c) system analyst
(d) DBA
53. _____ monitor and perform all activities related to database design, implementation, maintenance, and security.
- (a) End-users
(b) Application programmers
(c) System analysts
(d) DBAs
54. _____ identify the end-users' requirements and plan solutions.
- (a) End-users
(b) Application programmers
(c) System analysts
(d) DBA
55. _____ is a collection of interrelated data and a set of programs that enable multiple users to access, modify, share, and process it.
- (a) File (c) DBMS
(b) Database (d) Directory
56. There are _____ levels of data view.
- (a) 1 (c) 3
(b) 2 (d) 4
57. Which of the following is not a type of data view?
- (a) Tabular (c) External
(b) Internal (d) Physical
58. Which data view focuses on how data is actually stored?
- (a) Tabular (c) External
(b) Internal (d) Physical
59. What data is stored in database and the relationships that exist among that data is specified by which view?
- (a) Logical (c) External
(b) Internal (d) Physical
60. End users are interested in which data view?
- (a) Logical (c) External
(b) Internal (d) Physical

61. _____ view hides all complexities to deal with the user's view of the database.
(a) Logical (c) External
(b) Internal (d) Physical
62. _____ view enables customization of data according to user's needs.
(a) Logical (c) External
(b) Internal (d) Physical
63. _____ describes the structure of a database in a formal language supported by the DBMS.
(a) View (c) Architecture
(b) Schema (d) Template
64. There can be _____ number of external schema.
(a) 1 (c) 3
(b) 2 (d) any
65. There can be _____ number of logical and physical schema.
(a) 1 (c) 3
(b) 2 (d) any
66. There are _____ levels of database schema.
(a) 1 (c) 3
(b) 2 (d) any
67. _____ schema can be easily implemented without affecting the application programs that are using the data.
(a) Logical (c) External
(b) Internal (d) Physical
68. _____ schema is managed by the operating system under the direction of the DBMS.
(a) Logical (c) External
(b) Internal (d) Physical
69. _____ schema is often used by the application developers and programmers to develop applications.
(a) Logical (c) External
(b) Internal (d) Physical
70. _____ schema describes a part of the database as per the user's requirements and hides the rest of the database from that user.
(a) Logical (c) External
(b) Internal (d) Physical
71. In the three-schema architecture, each user group has its own _____ view.
(a) logical (c) external
(b) internal (d) physical
72. DBMS transforms the requests from _____ level.
(a) external, physical, logical
(b) external, logical, physical
(c) physical, logical, external
(d) external, logical, physical
73. The three-schema architecture ensures data _____.
(a) isolation (c) independence
(b) dependence (d) atomicity
74. _____ is the ability to modify the database schema at one level without affecting or changing the schema at the other levels.
(a) Data isolation
(b) Data dependence
(c) Data independence
(d) Data atomicity
75. The logical schema may be changed due to _____.
(a) addition of fields
(b) addition of constraints
(c) deletion of constraints
(d) all of these
76. _____ schema is changed to improve the performance of database.
(a) Logical (c) External
(b) Internal (d) Physical
77. _____ is the ability to change the internal schema without affecting the logical or external schema.
(a) Data dependence
(b) Physical data independence
(c) Logical data independence
(d) External data independence
78. Physical schema of the database may be changed due to _____.
(a) changing storage structure
(b) addition of indexes
(c) deletion of indexes
(d) all of these

79. Database _____ specifies integrity rules and type of operations that can be performed.
 (a) model (c) schema
 (b) view (d) all of these
80. Integrity rules ensure that the data is _____.
 (a) accurate (c) complete
 (b) current (d) all of these
81. _____ data model organizes data in a tree structure.
 (a) Network (c) Relational
 (b) Hierarchical (d) Object-oriented
82. _____ data model organizes data in a parent-child relationship.
 (a) Network (c) Relational
 (b) Hierarchical (d) Object-oriented
83. A parent can have _____ child/children in a hierarchical data model.
 (a) 1 (c) 3
 (b) 2 (d) n
84. If data is organized by state, within state by city, within city by zip code, then which data model will best suit the scenario?
 (a) Network (c) Relational
 (b) Hierarchical (d) Object-oriented
85. Data is represented by links in which data model(s)?
 (a) Network (c) Relational
 (b) Hierarchical (d) Both (b) and (c)
86. Which data model stores data in the form of graphs?
 (a) Network (c) Relational
 (b) Hierarchical (d) Both (b) and (c)
87. Child data element can have many parent data elements in which data model?
 (a) Network (c) Relational
 (b) Hierarchical (d) Both (b) and (c)
88. In relational model, each _____ stores a record.
 (a) row (c) field
 (b) column (d) file
89. In a relational database, each row represents a _____.
 (a) attribute (c) entity
 (b) value (d) all of these
90. Which data model is easy to expand?
 (a) Network (c) Relational
 (b) Hierarchical (d) Both (b) and (c)
91. Which type of data cannot be stored in a relational data model?
 (a) Multimedia (c) Spatial
 (b) Temporal (d) All of these
92. Which data model is used for CAD/CAM applications?
 (a) Network (c) Relational
 (b) Hierarchical (d) Object-oriented
93. Which data model is used for expert and multimedia systems?
 (a) Network (c) Relational
 (b) Hierarchical (d) Object-oriented
94. A set of all the objects which share the same attributes and methods is called a/an _____.
 (a) class (c) method
 (b) object (d) message
95. In a relational data model, rows are known as _____.
 (a) tuples (c) attributes
 (b) entities (d) table
96. When designing a relation, an employee will be treated as a/an _____.
 (a) attribute (c) relation
 (b) tuple (d) key
97. An entity is described using a/an _____.
 (a) attribute (c) relation
 (b) tuple (d) key
98. Which key uniquely identifies a row?
 (a) Primary key (c) Alternate key
 (b) Candidate key (d) All of these
99. _____ is an attribute or a combination of attributes that uniquely identifies an entity.
 (a) Primary key (c) Alternate key
 (b) Candidate key (d) All of these
100. Every table must have at least _____ candidate key.
 (a) 1 (c) 3
 (b) 2 (d) n
101. If a table has more than one candidate key, then one of them is chosen as the primary key and all the others are called _____ keys.
 (a) primary (c) alternate

- (b) candidate (d) composite
102. Primary key is also a _____ key.
 (a) candidate (c) both (a) and (b)
 (b) alternate (d) none of these
103. _____ is a key made up of two or more attributes within a table to uniquely identify a record.
 (a) Primary key (c) Alternate key
 (b) Candidate key (d) Composite key
104. _____ is used to cross-reference tables.
 (a) Primary key (c) Alternate key
 (b) Foreign key (d) Composite key
105. _____ acts as primary key of another table.
 (a) Foreign key (c) Alternate key
 (b) Candidate key (d) Composite key
106. _____ integrity checks are made when records are added, deleted, or updated in the table having the foreign key.
 (a) Data (c) Referential
 (b) Relational (d) Hierarchical
107. Name of table, attributes, primary key, foreign key, etc., are stored in a _____.
 (a) table (c) data dictionary
 (b) database (d) directory
108. Data dictionary is shared by _____.
 application(s).
 (a) 1 (c) 3
 (b) 2 (d) all
109. Data dictionary is usually hidden from the users so that it is not _____ by the users.
 (a) destroyed (c) accessed
 (b) modified (d) all of these
110. Documentation on database design process and information about data ownership is stored in the _____.
 (a) table (c) data dictionary
 (b) database (d) directory
111. The _____ statement of the SQL allows users to query or retrieve data from a table in the database.
 (a) select (c) retrieve
 (b) query (d) get
112. To select only specific data from the table, we use the _____ statement.
 (a) select - where (c) get - where
 (b) select - having (d) get - having
113. _____ is an information delivery system.
 (a) Table (c) File
 (b) Database (d) Data warehouse
114. In a data warehouse, data is stored by _____.
 (a) attributes (c) subjects
 (b) fields (d) applications
115. In a transactional system, data is stored by _____.
 (a) attributes (c) subjects
 (b) fields (d) applications
116. Before storing data in the data warehouse, it must be _____.
 (a) cleansed (c) filtered
 (b) transformed (d) both (a) and (b)
117. Standardizing data means making the _____ uniform.
 (a) field names (c) data layout
 (b) character code (d) all of these
118. Business transactions do not update the data in the data warehouse means data in the data warehouse is _____.
 (a) subject-oriented (c) time-variant
 (b) non-volatile (d) integrated
119. Data in the data warehouse is meant for analysis and decision-making. This requires that the data should be _____.
 (a) subject-oriented (c) time-variant
 (b) non-volatile (d) integrated
120. Time variant data allow users to _____.
 (a) analyse the past
 (b) forecast the future
 (c) relate information to the present
 (d) all of these
121. Data warehouse stores _____ data.
 (a) archived (c) transactional
 (b) current (d) highly detailed

122. _____ refers to using a variety of techniques to identify nuggets of information in the database.
 (a) Data warehousing (c) Data mining
 (b) SQL (d) IoT
123. Data mining is a/an _____ discovery process.
 (a) data (c) knowledge
 (b) information (d) results
124. Data _____ means summarizing general features of objects in a target class.
 (a) characterization
 (b) discrimination
 (c) association analysis
 (d) prediction
125. Data _____ produces rules that are used to compare general features of objects between two classes.
 (a) characterization
 (b) discrimination
 (c) association analysis
 (d) prediction
126. _____ technique studies the frequent item set.
 (a) Characterization
 (b) Discrimination
 (c) Association analysis
 (d) Prediction
127. _____ organizes data in given classes.
 (a) Characterization (c) Classification
 (b) Discrimination (d) Prediction
128. Forecast of missing numerical values is done using _____.
 (a) characterization (c) classification
 (b) discrimination (d) prediction
129. _____ is based on the principle of maximizing the similarity between objects in the same class and minimizing the similarity between objects of different classes.
 (a) Characterization (c) Classification
 (b) Clustering (d) Prediction
130. _____ is also known as unsupervised classification.
 (a) Characterization (c) Classification
 (b) Clustering (d) Prediction
131. Outlier analysis finds out the _____ in data.
 (a) outliers (c) surprises
 (b) exceptions (d) all of these
132. Study of trends in data is called _____.
 (a) evolution (c) classification
 (b) deviation analysis (d) prediction
133. Study of differences between measured values and expected values is called _____.
 (a) evolution (c) classification
 (b) deviation analysis (d) prediction
134. Data mining is best applied on _____.
 (a) table
 (b) database
 (c) data warehouse
 (d) transactional systems
135. Customer segmentation uses the _____ technique.
 (a) characterization (c) classification
 (b) clustering (d) prediction
136. Market basket analysis uses _____ technique.
 (a) characterization
 (b) discrimination
 (c) association analysis
 (d) prediction
137. Risk management uses _____ technique.
 (a) outlier analysis (c) classification
 (b) clustering (d) prediction
138. Which of the following is not a disadvantage of data mining?
 (a) Mining of information
 (b) Difficult to use
 (c) Privacy issues
 (d) None of these
139. Delinquency tracking uses _____ technique.
 (a) outlier analysis (c) classification
 (b) clustering (d) prediction
140. _____ is the process of collecting, organizing, and analyzing large sets of data (called big data) to discover patterns and other useful information.

- (a) Data mining (c) Both (a) and (b)
 (b) Big data analysis (d) None of these
141. Big data can be obtained from _____.
 (a) sensors (c) social media
 (b) video/audio (d) all of these
142. Hadoop is a tool for _____.
 (a) data warehousing
 (b) big data analysis
 (c) data mining
 (d) operational analysis
143. Zettabytes of data is analysed using _____.
 (a) data warehousing
 (b) big data analysis
 (c) data mining
 (d) operational analysis
144. Data warehouse can analyse _____ of data.
 (a) megabytes (c) gigabytes
 (b) terabytes (d) zettabytes
145. Cloud based analytics is used in _____.
 (a) data warehousing
 (b) big data analysis
 (c) data mining
 (d) operational analysis
146. Identify the latest breakthrough in discovering knowledge from data.
 (a) Data warehousing
 (b) Big data analysis
 (c) Data mining
 (d) Operational analysis

Answers to Multiple-choice Questions

- | | | | | | | | | | |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1. (a) | 2. (b) | 3. (c) | 4. (c) | 5. (a) | 6. (b) | 7. (a) | 8. (a) | 9. (b) | 10. (b) |
| 11. (d) | 12. (d) | 13. (b) | 14. (b) | 15. (d) | 16. (c) | 17. (b) | 18. (b) | 19. (b) | 20. (c) |
| 21. (b) | 22. (b) | 23. (b) | 24. (c) | 25. (d) | 26. (d) | 27. (d) | 28. (a) | 29. (a) | 30. (b) |
| 31. (c) | 32. (d) | 33. (a) | 34. (d) | 35. (c) | 36. (c) | 37. (b) | 38. (d) | 39. (c) | 40. (a) |
| 41. (d) | 42. (a) | 43. (b) | 44. (d) | 45. (a) | 46. (c) | 47. (c) | 48. (a) | 49. (b) | 50. (d) |
| 51. (d) | 52. (d) | 53. (d) | 54. (c) | 55. (c) | 56. (c) | 57. (a) | 58. (d) | 59. (a) | 60. (c) |
| 61. (c) | 62. (c) | 63. (b) | 64. (d) | 65. (a) | 66. (c) | 67. (d) | 68. (d) | 69. (a) | 70. (c) |
| 71. (c) | 72. (b) | 73. (c) | 74. (c) | 75. (d) | 76. (d) | 77. (b) | 78. (d) | 79. (a) | 80. (a) |
| 81. (b) | 82. (b) | 83. (d) | 84. (b) | 85. (d) | 86. (a) | 87. (a) | 88. (a) | 89. (c) | 90. (c) |
| 91. (d) | 92. (d) | 93. (d) | 94. (a) | 95. (a) | 96. (c) | 97. (a) | 98. (a) | 99. (b) | 100. (a) |
| 101. (c) | 102. (a) | 103. (d) | 104. (b) | 105. (a) | 106. (c) | 107. (c) | 108. (d) | 109. (d) | 110. (c) |
| 111. (a) | 112. (a) | 113. (d) | 114. (c) | 115. (d) | 116. (d) | 117. (d) | 118. (b) | 119. (c) | 120. (d) |
| 121. (a) | 122. (c) | 123. (c) | 124. (a) | 125. (b) | 126. (c) | 127. (c) | 128. (d) | 129. (d) | 130. (d) |
| 131. (d) | 132. (a) | 133. (b) | 134. (c) | 135. (b) | 136. (c) | 137. (a) | 138. (d) | 139. (a) | 140. (c) |
| 141. (d) | 142. (b) | 143. (b) | 144. (c) | 145. (b) | 146. (b) | | | | |

Internet and its Applications

Syllabus Mapping	Unit
Meaning of Internet, IP address [IPv4, IPv6], URL, domain name system, internet protocols—TCP/IP, UDP, FTP, TELNET [brief ideas only], HTML, DHTML and XML [concepts only], ethical hacking, cloud computing, mobile computing, Internet of things, ethical issues in social networking.	Module 1 Unit 3

3.1 INTRODUCTION

The Internet is a global network that connects billions of computers all over the world. It is a network of networks. The Internet links different organizations, academic institutions, government offices, and home users to share information among a large group of users.

Each computer on the Internet is called a host. To connect to the Internet, the user must gain access through a commercial Internet service provider (ISP). The Internet, sometimes known as the Net, allows the users to perform the following functions:

- Connect easily through ordinary personal computers (PCs) and local phone numbers to share a huge pool of information.
- Exchange emails with friends and colleagues as the email service has practically replaced the postal service for short written transactions and has undoubtedly become the most widely used application on the Net.
- Converse with other users on the Net where the conversation is text-based, voice-based, video-based, or a combination of all of them.
- Share important information in a timely manner.
- Access multimedia information that includes sound, photographic images, and video.
- Browse for information on websites using a web browser. The most popular browsers are Microsoft Internet Explorer, Netscape Navigator, Opera, Google Chrome, and Mozilla Firefox.

An additional feature of the Internet is that it lacks a central controlling authority. Although there are different governing boards that work to establish policies and standards, the Internet is bound by a few rules and answers no single organization.

3.1.1 History

The roots of the Internet can be traced back to 1969 when the Advanced Research Projects Agency (ARPA) of the US government formed the first network that was widely known as the Advanced Research Projects Agency Network (ARPANET).

This network was initially created to interconnect computers so that users in research organizations and universities could communicate with each other and share information.

However, it was in 1989 that the US government had lifted restrictions on the use of Internet and allowed it to be used for commercial purposes as well. Since then, the Internet has grown rapidly to become the world's largest network that connects thousands of networks, billions of computers, and hundreds of countries across the world.

3.2 INTERNET SERVICES

Today, the Internet has become a part of not only big organizations, universities, and offices but has also become the need of home users (like students) all over the world. In this era, life without the Internet has become unimaginable. Users access the Internet to avail services like email. In this section, we will read in detail about these services.

3.2.1 Electronic Mail

Electronic mail or email (Figure 3.1) refers to the means of transmission of messages electronically over communication networks. These messages may vary from notes entered from the keyboard to electronic files stored on the disk. Companies that are fully computerized make extensive use of emails because it is fast, flexible, and reliable.

Email is one of the most widely used services on the Internet. Anyone with an email account can send an electronic mail (like a letter) to any other person who also has an email account (provided the email address of the recipient is known). Usually, the structure of the email can be given as *username@domain name*.

For example, if a user has created an email account on Gmail, then the email address is *username@gmail.com* where *username* is the Gmail screen name.

Using an email is rather straightforward. If you have an email account, you can just click on the option that says something similar to Compose e-mail. Thereafter, you will be prompted to enter the following fields.

To In this field, the email address of the recipient will be written.

Subject In this field, you can type anything illustrating the content of the message. However, the length of this field is limited.

Body In this field, the content of the message is written. Most email systems include a rudimentary text editor for composing messages but users may also edit messages using any other editor of their choice.

CC and BCC Filling up these fields is optional. However, similar to the *To* field, the email address of the recipient is also entered in these. While *CC* stands for carbon copy, *BCC*, on the other hand, means blind carbon copy. When you wish to send the same email message to multiple recipients, add the extra email addresses in the *CC* field, separating each address with a comma. *BCC* works just like *CC*, except that the email addresses in *BCC* are not revealed to the other recipients. For example, if you send an email *To: Goransh* and *BCC: Radhika*, then *Radhika* will see *Goransh's* email address, but *Goransh* will not see *Radhika's* email address.

Figure 3.1 Email

Moreover, there are also options for attachments and forwards. These options enable users to add files to the email. The forward option is used to forward (make a copy) a message received from someone and mail it to someone else.

The sent messages are stored in electronic mailboxes until the recipient fetches them. After reading the message, the receiver may save it, delete it, forward it to someone else, or reply to it. On the whole, email is a good alternative to the traditional paper-based letters. It is much faster and easier to use as there is no need to buy envelopes and stamps and involves no queuing in the post office.

3.2.2 File Transfer Protocol

File transfer protocol (FTP) is one of the oldest applications of the Internet for transferring files from one computer to another like from a user's PC to a web server. FTP is discussed in more detail in section 3.7.4.

3.2.3 Chatting

Internet chatting is a very popular service of the Internet that allows two or more online users to come together to talk using an instant messenger. Chatting helps users to stay connected with concerned people in business or family who live many miles away. Though chatting can be fun and entertaining, users must take protective measures to avoid Internet stalkers and predators.

However, to chat with users on the Internet, every user must have an account with a username and password to enter the website. Chatting involves the exchange of typed-in messages between a group of users who take part from anywhere on the Internet. The chatting program also enables users to arrange a private chat between two parties who met initially in a group chat.

Nowadays, even business organizations are using chatting services to host online business meetings, answer any queries of customers, or provide them with online support and assistance. Whether a business chat or a personal chat, it can be ongoing or scheduled for a particular duration. Most chats are focused on a particular topic of interest and some involve guest experts or celebrities to talk to other online members who want to join the chat.

3.2.4 Internet Conferencing

The evolution of the Internet has changed the way in which business houses arrange conferences. Nowadays, organizations are increasingly switching to Internet conferencing to reduce the extra costs involved in travelling and making telephone calls, thereby resulting in better time management and enhanced productivity.

Internet conferencing is quite similar to traditional teleconferencing. The difference is that during an Internet conference call, participants will sit at their respective offices while being connected to each other through the Internet. Many Internet conferencing software allow as many as 25 online participants to participate in a web meeting.

Today, there are numerous software available on the Internet that can be easily downloaded to host a conference call at little or no cost. To start an Internet conference call, users need a computer with an Internet connection, related software, a webcam, and a microphone (to see and hear what the remote or distant participants are saying). During the conference, users can share information, files, video, and audio clips.

Internet conferencing has taken virtual meetings (because users are not present at the same place) to a new level. This service is increasingly being used to collaborate with various teams in a company to

deliver presentations, to host small-scale or large-scale seminars, to connect with technical support, and so on. The advantages of Internet conferencing are multiple and aid small and established businesses alike.

3.2.5 Electronic Newspaper

An online newspaper, also known as a web newspaper or an electronic newspaper, is a newspaper on the Internet, which is either published separately or as an online version of a printed periodical.

With online newspapers, users can read the full coverage of breaking news in a timely manner. The credibility and strong brand recognition of well-established newspapers, close relationships they have with advertisers, and savings in overhead costs have led to a shift from the traditional printing process to online coverage.

Online newspapers are almost similar to the hard-copy newspapers and have the same legal boundaries such as laws regarding privacy and copyright issues. Moreover, news reporters are being taught to shoot videos and to write in the succinct manner necessary for news pages on the web.

3.2.6 World Wide Web

The technical definition of the World Wide Web (WWW) can be given as 'all the resources and users on the Internet that are using the hypertext transfer protocol (HTTP)'. On the web, all the documents are formatted in a special markup language called hypertext markup language (HTML) that supports links to other documents, graphics, audio, and/or video files. This feature enables users to jump from one document to another simply by clicking on hot spots. You must have noted that when you position your cursor on a hotspot also known as hyperlink, the cursor changes to a hand-shaped figure. When you click on the hyperlink, you are taken to another part of the information.

In simple terminology, the WWW is a part of the Internet that allows easy navigation through the use of GUIs and hypertext links between different addresses.

The WWW was created in 1989 by Tim Berners-Lee. The Web, in simple terms, is the user part of the Internet. Novice or professional users make use of the Web to communicate and access information for business and recreational purposes. There are several applications called web browsers that make it easy to access the web. Some popular web browsers are Mozilla Firefox, Opera, Google Chrome, Netscape Navigator, and Microsoft's Internet Explorer.

Many a time, we think that the Internet and the WWW are the same, but this notion is not correct. The Internet and the Web work together. While the Internet provides the underlying structure, the Web on the other hand, utilizes that structure to offer content, documents, multimedia, etc. For example, the Internet is like the highway, and the WWW is like a truck that uses that highway to get from one place to another.



Hypertext is a method of instant cross-referencing. On websites, certain words or phrases appear in text of a different colour than the rest; and are often underlined. When such words or phrases are selected, users are transferred to the site or page that is relevant to this word or phrase. Sometimes, there are buttons, images, or portions of images that can be clicked.

3.2.7 Online Shopping

Online shopping refers to buying goods and/or services from merchants who sell on the Internet. The popularity of the World Wide Web gave an excellent opportunity to merchants to sell their products to

people who surf the Internet. Shoppers can now visit web stores 24x7 as per their convenience while sitting in their homes, offices, or even while travelling.

Consumers buy a variety of products from online stores ranging from books, clothing, household appliances, toys, hardware, software, to health insurance policies.

Nowadays, there is a rise in the number of online shoppers. This is mainly because when a customer shops at a brick-and-mortar store, he has to drive to the store, find a parking place, and walk throughout the store until he locates the products that are needed. Even after selecting all the products, the customer has to wait in long queues to get the billing done. In striking contrast to this, online shopping is very convenient. With online shopping, a shopper logs onto the Internet, visits the store's website, and selects the products. Once all the products have been selected, the user can click the option to calculate the invoice and pay for the purchase using his debit or credit card.

Although online shopping seems to be very convenient, many people still choose to avoid it. This can be due to several reasons. First, some people relish the experience of shopping in a mall. Moreover, they like to touch the merchandise, try on clothing, and be around other people. However, with online shopping, shoppers cannot touch the products or have any social interaction.

Second, people fear that their credit card information will be compromised. Since it is necessary to provide credit card information when purchasing products online, they fear that their personal information may be susceptible to theft.

Third, consumers often feel that the products they purchase are not accurately portrayed in the website's picture. The picture of an item may appear in a particular way, but the actual item may look completely different—perhaps of lesser quality. Many a time, it is also impossible to try on apparel when conducting online shopping.

3.3 INTERNET PROTOCOL ADDRESS

When you have to send a letter to your friend, you always mention the residence address of your friend so that it can reach her. This residence address is unique. If you want to talk to your friend, you need to dial her telephone number, which is again unique. Similarly, if you have to send a message over the Internet, you need some addressing mechanism so that the message can reach the correct destination.

Internet protocol address or IP address is, therefore, a unique address allotted to computing devices such as computers, routers, printers, scanners, modems, smartphones, tablets, and so on that are connected with the Internet. This address facilitates unique identification of devices for communication to take place.

Features

The following are the features of IP address:

- An IP address is divided into four parts where each part is separated from the other using a dot.
- Each part of the address contains a number ranging from 0–255.
- For example, 79.121.10.190 is a valid IP address.
- Without IP addresses, sending and receiving data over the Internet would be impossible.

3.3.1 Types of IP Addresses

There are two categories of IP addresses—static or dynamic and public or private.

Static and
change an
Many a tim
remains v
lease. Sim
automatic
the users.

Table 3.

Static

This IP

A com
it acc

It is a
betwe

This
the c
com

Ther

It is
Inte

Em
sta

It i
IP,
to
on

It
m

A

Pr
en

(N
T

re
o

n

Static and dynamic Internet protocol addresses As the name suggests, static IP addresses never change and dynamic IP addresses keep changing dynamically whenever users log on to the network. Many a time, dynamic IP addresses are issued using a leasing system. Therefore, the allocated IP address remains valid for a limited time. When the lease expires, the computer automatically requests a new lease. Similarly, when there is an IP address conflict, a request for another IP address is sent to the ISP automatically. The entire process of requesting IP addresses is automated and is therefore hidden from the users. Table 3.2 shows the differences between static and dynamic IP addresses.

Table 3.2 Differences between static and dynamic IP address

Static IP address	Dynamic IP address
This IP address is permanent.	This IP address is temporary and changes dynamically.
A computer retains its static address every time it accesses the Internet.	A computer is allotted a new dynamic IP address every time it accesses the Internet.
It is a reliable way to facilitate communication between remote devices.	It is not a reliable way to facilitate communication between remote devices.
This address reveals technical information about the continent, country, and city in which the computer is located.	Dynamic IP address does not reveal any such detail.
There are limited static addresses.	It can support a large number of users who do not require the same IP address always.
It is allotted to devices that connect to the Internet using a broadband connection.	It is allotted to devices using a dial-up connection to connect to the Internet.
Email servers and other web servers must have a static IP address.	It is not suitable for servers.
It is preferable for applications such as voice over IP, online gaming, and other applications that need to locate and connect to a particular computer on the Internet.	It is preferable for applications that work fine with temporary and one-time IP addresses.
It is less safe and requires extra security mechanism.	It is safe to use.
A user can configure his static IP address himself.	Dynamic IP address is allotted by the ISP server. This allocation is transparent from the user.

Public and private Internet protocol addresses Certain computers such as web servers and mail servers need to maintain a unique global IP address that is registered with the network information centre (NIC) to avoid address conflicts. This address is publically known and used by users all over the world. Therefore, such an IP address is called public IP address.

In striking contrast with public IP addresses, private IP addresses are allocated to devices that do not require public access. These devices are on the network but need to be uniquely identified only within an organization. For example, a network printer is assigned a private IP address to ensure that people from rest of the world are not able to connect with it.



The NIC has reserved certain addresses for private use which organizations can use to allow them to their devices that need not be identified over the Internet.

3.4 DOMAIN NAME SYSTEM

When you want to talk to a friend, you do not type his number. You may be having several friends and memorizing everyone's number is just not possible. Therefore, you save all important phone numbers along with their names in your phonebook. To connect with your friend, instead of dialling his 10-digit cell number, you just search for his name in the phonebook and click the *Call* button. Although you use the name, your call is not connected based on name. The name is converted into a number which is then used to establish the connection.

Coming back to the Internet, we have seen that every device has a unique IP address. To connect with a particular device you need to specify its address. However, we do not really type the IP address. For example, if we want to connect to google.com, we just type www.google.com. Then where is the IP address and how are we able to access the website? The answer to this question is the domain name system (DNS). Similar to the phonebook service, the Internet has a corresponding DNS service that translates domain names into IP addresses (for example, www.google.com into 74.125.224.72). This means that every time we use the Internet, we always use the DNS.

The DNS system works as a network of DNS servers. As maintaining a central database of all the computers on the Internet along with their names and IP address is quite impractical, the DNS distributes the responsibility of storing domain names and their corresponding IP addresses to authoritative name servers. These name servers are responsible for the domain they support. The authoritative name servers may even delegate authority to other sub-domain servers. Besides providing speedy mapping, this authority delegation process ensures distributed and fault-tolerant service to Internet users.

In such a networked DNS environment, if one DNS server does not know the IP address of a particular domain name, it asks another server for the same. The process is repeated until a proper match between IP address and domain name is found. This concept is shown in Figure 3.3.

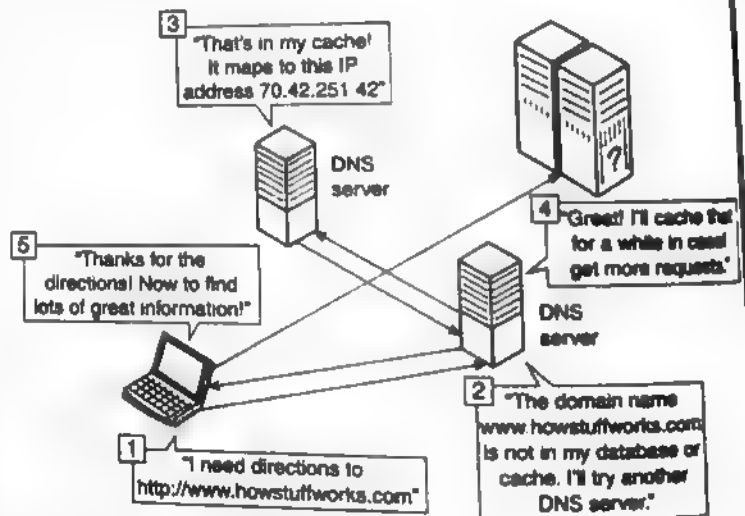


Figure 3.3 Working of DNS



DNS is a service that automatically converts domain names into IP addresses.

Points to Remember

Here are some key points about DNS:

- It assumes that IP addresses are assigned statically and will not change.

- It supports caching of requests.
- ISPs maintain their own DNS servers to resolve name to IP address mapping.
- Since DNS translation causes additional overhead when accessing any website on the Internet, ISPs cache DNS mapping in their DNS server so that that can automatically direct subsequent requests to the appropriate IP address.

Some commonly used domains are as follows:

gov Government agencies	com Commercial business
edu Educational institutions	net Network organizations
org Non-profit organizations	int International organizations
mil Military	

Some country domains are as follows:

ca Canada	in India
th Thailand	us United States of America
fr France	uk United Kingdom
jp Japan	

Other domain names include .museum (for museums), .info (informational websites), .name (personal websites), .pro (for professionals), .aero (for aeronautical companies), .coops (for co-operative organizations), .jobs (for job posting), .mobi (for mobile communication networks), etc.

Like our full names in which the general name or surname comes on the right and our specific name comes on the left, domain names are also organized from right to left, with general domains to the right, and specific domains to the left. For example, in the domain name www.google.com, there are three domain names, each separated by a dot. Here, .com is a general domain and google is a sub-domain, and www is a sub-domain prefix for the World Wide Web.

3.5 UNIFORM/UNIVERSAL RESOURCE LOCATOR

A uniform resource locator (URL) specifies the unique address for a file that is accessible on the Internet. It is provided by the user in the address bar. For example, when you type www.google.com, after pressing the Enter key, there is a long sequence of characters in the address bar. This is the URL. This means that to access any page on the Internet, we need to provide its URL.

The file on the Internet that we want to access can be a web page, an audio file, video file, or image with extensions .htm, .php, .mp4, .avi, .jpg, .bmp, .gif, .asp, .cgi, .xml, etc.

The syntax for a URL is as follows:

Protocol://domain-name/path

where protocol specifies the name of the protocol to be used to access the file resource. Commonly used protocols are http, https, ftp, telnet, news, gopher, mailto, etc. This field specifies how to connect.

Domain name identifies the name of the website. This means that the domain field identifies where to connect.

Path is a hierarchical description that indicates the location of the file. It indicates to the web server what to connect.

For example, when we just write <http://www.google.com>, http is the protocol, www.google.com is the domain name, and by default, the home page which is saved as index.htm is displayed to the user. Refer to Figure 3.4 which shows another sample URL.

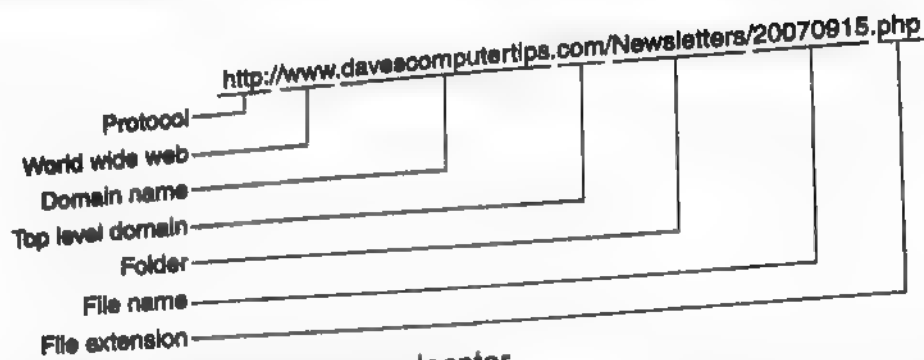


Figure 3.4 Uniform resource locator

If we provide the URL as, `http://www.example.com/Student/ABC.TXT`, then `http` is used to fetch the file `ABC.TXT` from `Student` directory stored in the computer on which the website `www.example.com` is hosted.

There are basically two types of URLs as shown in Figure 3.5. While an absolute URL specifies the complete URL containing all three fields (protocol, domain, and path), relative URLs, on the other hand, contain only the one field which is the domain name.

Many a time, you must have observed a complex URL as the one given here, especially when you log in to your email account or search for a string on google.

`http://www.google.com/cgi-bin/search.cgi?q=computer%20fundamentals`

Although it seems complex, it is actually very simple to interpret. In the query, `http` is in the protocol, `www.google.com` is the domain, and `search.cgi` is a file in the `cgi-bin` directory. Anything following the question mark (?) in a URL is a pair of variable(s) and its value(s). In the URL, `q=computer%20fundamentals` means that `q` is a variable name, and `computer%20fundamentals` is the value of `q`. Since blank spaces are not allowed in a URL, blank space has been written as `%20`. Spaces can also be written as a `+` (plus) sign. In the query, the user is trying to search computer fundamentals on Google.

These values are sent by the user's computer to Google's server. Google will find relevant pages and will display the result on the user's screen. Similarly, when we log in to our email account, we supply two values to the server—username and password. In such a situation, multiple variables are separated with an ampersand (&) sign as shown here:

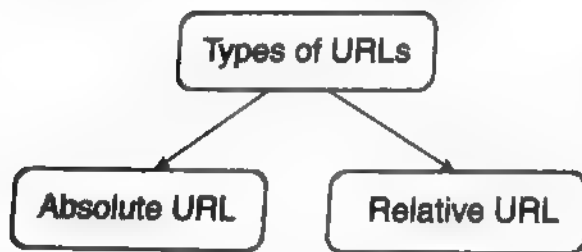


Figure 3.5 Types of URLs

`http://www.gmail.com/config/passwd.cgi?u=erree&p=s2ejmd3`

In the URL, there are two different variables—`u` with value `erree` and `p` with value `s2ejmd3`.



A domain name is not the same as URL because it is just a small part of the entire URL.

3.6 INTERNET PROTOCOLS

Each layer in the open system interconnection (OSI) model supports one or more protocols to exchange data between the source and destination machines. In this section, we will read about vital protocols that form the backbone of data communication over computer networks.

3.6.1 Internet Protocol

The term 'protocol' means a set of rules that must be followed to facilitate communication among diverse devices on the Internet. Internet protocol (IP) is basically a standard networking software which is pre-installed in your computer to enable you to interact with any computer in any country irrespective of the device and its manufacturer. This means that the same IP software is installed in a laptop, desktop, printer, or any other device that needs to be connected with the Internet.

In order to transfer data over the Internet, all devices use the TCP/IP technology. This technology, in turn, comprises various protocols such as TCP, IP, FTP, simple mail transfer protocol (SMTP), and so on. You may think of TCP/IP technology as a team of robot soldiers who are supposed to receive, handle, and transfer the data to the correct destination device.

Currently, two versions of IP are being used—IPv4 and IPv6 where v stands for version.

IP version 4 Currently, IPv4 is being widely used by most network devices. It is a 32-bit number (in binary) and can support a maximum of 2^{32} or 4.3 billion devices on the Internet. Addresses in IPv4 consist of 32 binary bits. The 32 bits are divided into four groups of eight bits where each group is separated by a dot. For example, 216.27.61.137 is an IP address written in decimal notation for better readability and understandability. The same address, when written in binary, can be given as 11011000.00011011.00111101.10001001. Figure 3.6 illustrates another example of an IP address using IPv4.

Although simple, the problem with version 4 is that as more and more devices are being added to the Internet every year, IPv4 addresses are getting exhausted.

IP version 6 The newer version of IP is IPv6 and is slowly replacing IPv4. IPv6 is not more advanced than IPv4 but has many new features. Since an address in version 6 is 128-bits long, it can support 2^{128} devices on the Internet or approximately, 340, 282, 366, 920, 938, 463, 463, 374, 607, 431, 768, 211, 456 devices. Due to their large size, the address in IPv6 is specified in hexadecimal separated by colons. For example, 1124:1:0:C:0:42:0:512C is a valid IP address specified in version 6. Presently, IPv4 and IPv6 addresses exist but soon IPv6 will take over version 4.

An IPv4 address (dotted-decimal notation)

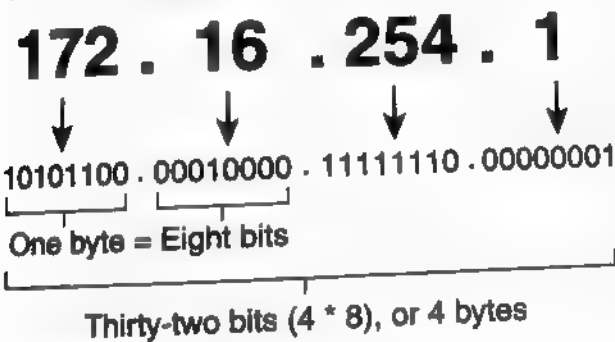


Figure 3.6 IP address



IPv5 is an experimental protocol for Unix-based systems and was never released to be used by the general public. All odd-numbered versions are developed for experimental purposes.

Table 3.3 highlights the differences between two versions of the IP.

Table 3.3 Differences between IPv4 and IPv6

IPv4	IPv6
Provides a 32-bit address	Provides a 128-bit address
Expressed in decimal format	Expressed in hexadecimal format
Can support 4.3 billion devices	Can support an infinitely large number of devices

Getting obsolete

Comparatively less compatible

Comparatively small-sized data can be sent

Comparatively less secure

Technologically more advanced with new features

More compatible for mobile networks

Bigger-sized data can be sent

More secure than IPv4

3.6.2 Transmission Control Protocol

Transmission control protocol (TCP) works at the transport layer. It is used in conjunction with the IP protocol which works on the network layer and is responsible for sending data packets between sender and receiver devices. The working of IP can be compared with the postal system which allows users to address a letter and submit it to the postal system that has no direct link between the sender and the receiver. The main responsibility of the TCP is, therefore, to establish a connection between the sender and the receiver so that reliable and error-free data transmission can take place. The connection established by TCP exists until both the devices have finished their data transmission.

Another important responsibility of TCP is segmentation and reassembly as shown in Figure 3.7. At the sender's site, TCP breaks the message into several smaller segments and at the receiver's site, TCP reassembles the fragmented data to form the complete message. Since each data packet of the same message may reach the receiver through different routes, it may happen that packet 3 arrives before packet 1. Therefore, the TCP at the receiver's site rearranges the packet in the correct sequence and then assembles them to form the complete message.

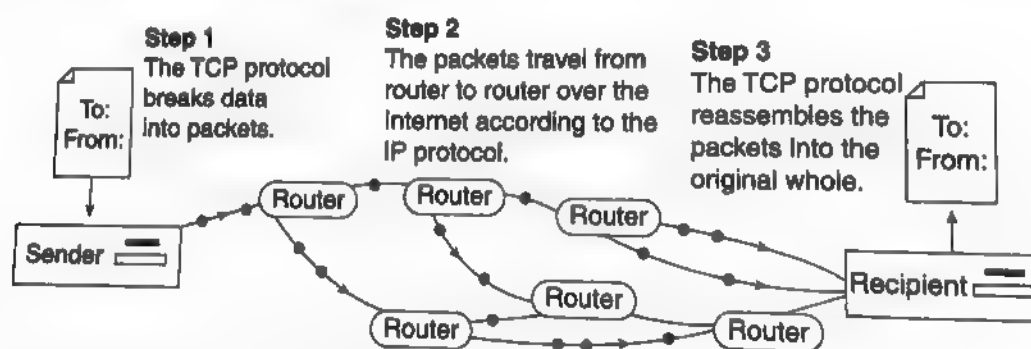


Figure 3.7 How TCP/IP works

3.6.3 User Datagram Protocol

Like TCP, the user datagram protocol (UDP) is a communications protocol that works on the transport layer (layer 4). The combination IP and UDP is known as UDP/IP. UDP is an alternative to TCP. While some applications prefer to use TCP, others make use of UDP. UDP offers limited service as compared to TCP. Unlike TCP, UDP does not break the message into smaller segments; therefore, it need not even reassemble or sequence them. If some erroneous data is received, UDP does nothing to ask the sender to resend it. For these reasons, many network applications such as trivial

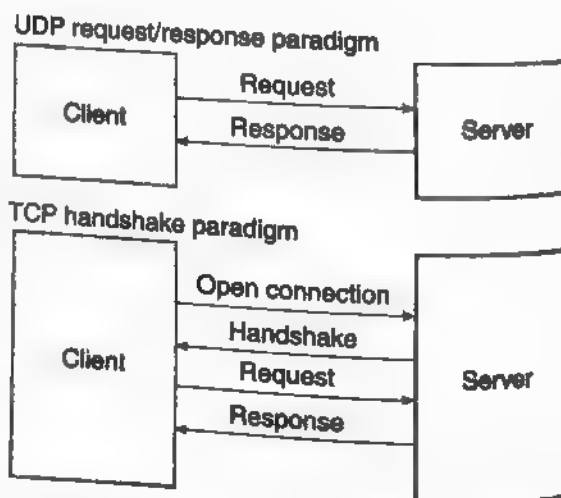


Figure 3.8 How TCP and UDP work

file transfer protocol (TFTP), videoconferencing, online computer games, watching videos online, and listening to audio online may want to save processing time because for very small data, exchanges prefer to use UDP rather than TCP. Table 3.4 and Figure 3.8 highlight the differences between TCP and UDP.

Table 3.4 Differences between transmission control protocol and user datagram protocol

TCP	UDP
Breaks the message into smaller segments at the sender's end	Does not break the message into smaller segments
Reassembles the segments at the receiver	No need to reassemble since the message was never broken
Preferable when data to be exchanged is large	Preferable for small data exchanges
Slower as compared to UDP	Faster since it provides limited services
Asks sender to resend erroneous data	Does not ask sender to resend erroneous data
Supports error checking and error correction mechanisms and ensures a reliable data transfer	Does nothing to correct errors and supports unreliable data transfer
Is a connection-oriented protocol	Is a connection-less protocol
Before data exchange, handshaking signals exchanged to establish connection	Data exchange is not preceded by any handshaking signals
Sends acknowledgement for data that has been received	No acknowledgments are sent

3.6.4 File Transfer Protocol

FTP is the preferred method of exchanging files because it is faster and reliable. It is usually used to perform the following functions:

- Enable users to share files, computer programs, and/or data.
- Enhance the use of remote computers.
- Provide authentication and security to stored files and programs.
- Transfer data reliably and efficiently.

Generally, users use FTP to transfer web page files from their computers to the server so that the pages can be available for everyone on the Internet. Using FTP, users can easily update (delete, rename, move, and copy) files on a server. However, it is also used to download programs and other files to the user's computer from other servers. The FTP program can be invoked either by typing the FTP commands on a simple command line interface (like the MS-DOS Prompt window) or with a commercial program that offers a graphical user interface (GUI). Even the Web browser can make FTP requests to download the selected programs.

In FTP terminology, the following rules are followed:

- The machine on which the file exists is called the server and the machine that requests that file is called the client.
- Copying files from a client to a server is called uploading, whereas transferring files from a server to a client is called downloading.
- The term 'get' refers to receiving files from the server and 'put' refers to sending files (synonymous with download and upload respectively).

How to Use FTP

Nowadays, using the FTP is as simple as using Windows Explorer. FTP is widely being used to upload web pages on the Internet. Users create web page files on their computers and transfer them to the ISP's web server by using FTP.

To enable this, an FTP client software is already installed in the user's computer. This software allows users to log in to a remote computer. The software has two window panes; the one on the left displays all files and directories in the user's computer and the pane on the right displays all files and folders present in the remote computer. Files can be easily transferred by simply dragging and dropping from one pane to another.

Steps for Using FTP

The following are the steps to be followed to use FTP:

- Open the FTP client software installed in computer. You can also use the FTP through your web browser by just typing the URL of the FTP server. However, web browsers are slower and less reliable than dedicated FTP clients.
- Enter the name of the FTP host (e.g., `ftp.microsoft.com`).
- Enter your username and password. If it is an anonymous FTP server like Cyberduck for Mac or WinSCP for Windows, then type anonymous as username and your email_id as your password. Usually, all publically available files are accessed using anonymous FTP server.

Using FTP through Command Line Instructions

While GUI-enabled FTP clients need to be downloaded from the Internet, users can also use FTP through command line instructions in Windows, Mac OS X, and Linux by typing `ftp ftp.microsoft.com`. Basic FTP support is provided by all computers.

Provide your username and password. In case of an anonymous FTP site, enter anonymous as username and email address as password.



FTP also allows users to delete, rename, move, and copy files on the server.

Modes of File Transfer

FTP can transfer files in the following three modes.

Stream mode It is that mode in which it transfers files as a continuous stream with no intervention or processing of information into different formats.

Block mode In this mode, the data to be transferred is divided into blocks.

Compressed mode In this mode, FTP compresses the files by encoding them.

In Figure 3.9, FTP uses two connections for data exchange—one for sending commands and the other for sending or receiving data.

Active and passive connection mode FTP supports two modes of connection—active and passive.

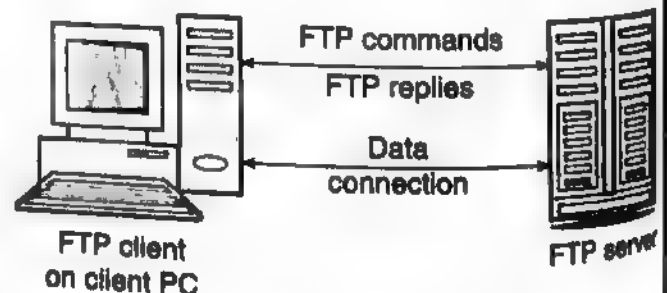


Figure 3.9 Two connections in FTP

In active mode, the client initiates the connection. Once the server gets connected with the client, the data can be exchanged.

In passive mode, the server is always waiting for any request from the client. Most FTP clients use passive connection mode by default.

3.6.5 Terminal Emulation

Terminal emulation or Telnet is a text-based protocol that uses the underlying TCP/IP technology for accessing a remote computer's (called host) data and application programs. Users can use the Telnet client software to connect with the Telnet server (or the remote host) as shown in Figure 3.10. Once the Telnet client establishes a connection to the remote host, the client becomes a virtual terminal and can communicate with the remote host from his computer.

Features

The following are the features of Telnet.

Bidirectional It is used to send and receive information. This means that data travels in both directions.

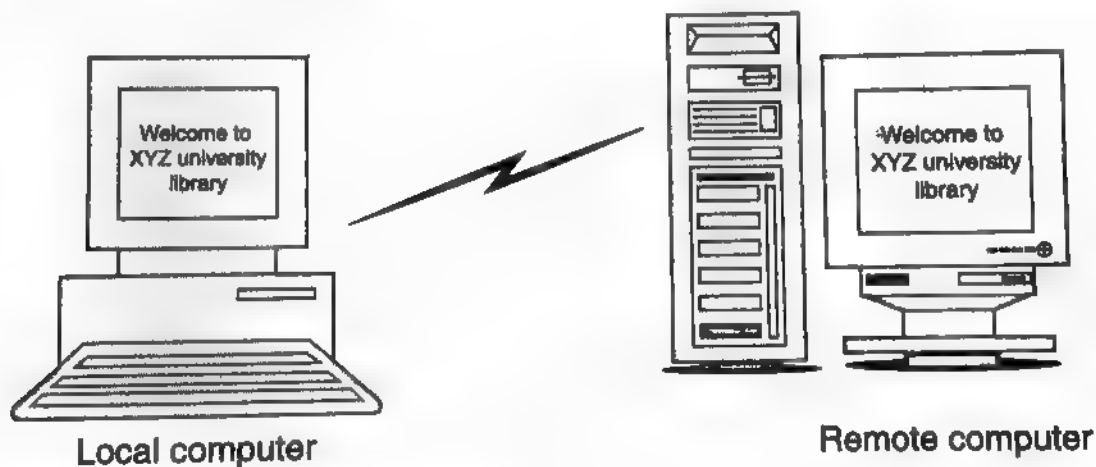


Figure 3.10 Telnet connection

Interactive It accepts a command from the user, executes it on the remote computer, and displays the result on the user's screen.

Text-oriented It is a text-based protocol that lacks usage of GUI.

Versatile It can be used with any device that uses either the phone connection or the Internet connection.

How to Use Telnet

Telnet client programs are available for all major operating systems. To use these clients, go to command prompt by clicking Start->Run and type the following:

```
telnet the.libraryat.whatis.edu
```

where Telnet is the name of the protocol and the.libraryat.whatis.edu is the name of the host. You can also specify the IP address of the host instead of its name. The syntax of using the Telnet command is as follows:

telnet host port (where port is optional)

Next, enter your username and password to log in to the remote computer.

Once the connection is established, users can enter commands which will be executed by the remote computer.

Uses

Telnet is used in the following areas:

- It is used to access specific applications or data located at a remote computer.
- It allows users to log in as users of the remote host.
- Even a dumb terminal can use the applications hosted on the world's most powerful computer.
- It enables users to control the server and communicate with it over the network.
- It enables research scholars and professors to log in to the university's computer from any terminal.
- It allows users to connect them to databases, library catalogues, and other information resources around the world.



Telnet has now evolved as secure shell (SSH) and is used by network administrators to manage Unix and Linux computers from a remote computer.

Shortcomings

The drawbacks of Telnet are as follows:

- It is a text-based computer protocol and lacks a fancy screen with images, animation, and hyperlinks.
- It is insecure because it transfers all data in clear text.
- Users cannot transfer files using Telnet.

3.6.6 Hypertext Markup Language

Hypertext markup language (HTML) is rightly said to be the mother tongue of web browsers. It is a standard language understood by all web browsers. HTML consists of a set of markup symbols or codes (commonly known as tags) inserted in a file to be displayed by the web browser such as Google Chrome, Mozilla Firefox, Netscape Navigator, Internet Explorer, and so on. In order to understand HTML, let us first break up these words and try to understand their individual meaning.

Hypertext The term 'hypertext' is opposite to linear. It is a method by which we browse the web. By clicking on a hypertext (where the mouse pointer turns into a finger), another text on the same web page or a different web page is displayed.

Markup Markups are the building blocks of HTML. A markup is a code embedded in a file that instructs the web browser how to display the page. Tags are displayed within angular brackets (e.g., <bold>).

Language HTML is a means of communicating instructions for formatting a hypertext document. It consists of codes and syntax like any other language.

HTML consists of numerous predefined tags to describe page content. Each opening tag (<...>) has a corresponding closing tag (</...>). The content to be displayed is written in between the tags. The tags instruct the web browser to display the content as specified by the tag. For example, the statement in HTML code,


```
<bold> Hello, World!!! </bold>
```

will be displayed as

Hello, World!!!

on the screen. From the example, we can observe that the tags themselves do not appear on the web page when you view it through a web browser. The web browser reads the tag, interprets it, and applies its effect while displaying the page. Some tags are also used to insert images, text, videos, tables, and forms in a web page.

Uses

The following are the uses of HTML:

- It is used for designing web pages.
- With HTML, web page designers can embed text, images, audio, video, and interactive forms in a web page.
- It allows designers to embed scripts (such as JavaScript or VBScript) within a web page to design interactive pages.

Although HTML is easy to learn and is improved every year, presently, web pages are rarely designed using only HTML.



You can view the HTML code of a web page by right clicking on a web page and selecting the 'View Source' option from the pop-up menu.

3.6.7 Dynamic HTML

Dynamic HTML (DHTML) is a combination of technologies—HTML, Java Script, and Cascading Style Sheets that is used to create dynamic, interactive, and animated web pages. DHTML is purposely used to enhance the user's experience. It processes data at the client's machine without sending it to the server for processing. Unlike HTML, DHTML is not a language. It is just a technology that allows web documents to look and act like desktop applications or multimedia productions.

Uses

The uses of DHTML are as follows:

- It enables a web designer to control the display and positioning of HTML tags in a browser window.
- It allows users to create dynamic, interactive, and animated web pages.
- It facilitates web pages to change at any time, without returning to the web server first.
- It allows users to change the looks and functionality of a web page after the page has been fully loaded.
- It enables users to add effects to their pages to make them more presentable and attractive.
- It allows users to use DHTML to embed animated text and images in their documents; for example, users can move a text or image from a starting point to an ending point, either by following a pre-determined path or one chosen by the user.
- It allows users to display information that changes dynamically and rapidly; for example, the latest news, stock quotes, etc.
- It designs a form to capture user input, verifies if all the fields have been filled, and responds to users.

- It incorporates rollover buttons or drop-down menus.
- It includes dynamic menus in a web page.
- It applies dynamic styles; for example, changing the background colour, font colour, font size, or content dynamically.
- It is used for creating games.
- It uses dynamic effects to enhance the functionality of a web page.

Concerns

The following are some concerns regarding DHTML:

- Excessive use of effects may frustrate the users. For example, a website menu with a lot of animations can confuse user navigation.
- Developing and debugging web pages with dynamic effects can be difficult.
- DHTML scripts may not work correctly in some web browsers. Therefore, web page designers must test the web page in all leading web browsers.
- Web pages may not be displayed correctly when they are viewed using different screen sizes.

Therefore, web page designers prefer to avoid complex DHTML effects and use simple routines that enhance user experience.

3.6.8 Extensible Markup Language

Like HTML, extensible markup language (XML) is a markup language that defines a set of rules for encoding documents in a format that is readable by humans as well as computers. XML is used to share information in a consistent way. It has been specifically designed for web documents. Like HTML, the basic building block of an XML document is a tag.

Features

Some of the features of XML are as follows:

- It is a flexible markup language as it allows web page designers to create their own customized tags.
- It is a text-based data format.
- It is easy to use and specifies a general format for describing data that enhances its usability over the Internet.
- It is widely used for representing arbitrary data structures.
- It is designed to be self-descriptive.
- It supports the use of nested tags to represent hierarchical data.

Table 3.5 summarizes the differences between HTML and XML.

Table 3.5 Differences between HTML and XML

HTML	XML
It is designed to display data.	It is designed to describe data.
HTML tags are predefined.	XML tags are not predefined.
It focuses on how data looks.	It focuses on what data is.
It is about displaying information.	It is about carrying information.

It does not need any extra piece of software to display data.

It just describes data. It needs another piece of software to display that data.

Let us take an example to see how data is described in XML with customized tags. In the example, Aditya sends a message to Sarthak.

```
<message>
<to>Sarthak</to>
<from>Aditya</from>
<heading>Reunion</heading>
<body>This Friday we have an alumni meet in our school. Let's meet there.
</body>
</message>
```

As we can see from the example, the message is self-descriptive. It has customized tags that represent information about sender, receiver, heading, and body message. These tags are not defined in any XML standard and are invented by the author of the XML document.

The example justifies that XML is self-describing or self-defining. Since the structure of the data is embedded with the data, it can be easily understood. It is, therefore, used to share information in a consistent way.

Final Note

XML is not a replacement for HTML; rather, it complements HTML. In web applications, XML and HTML are used together. While XML is used to describe data, HTML is used to format and display that data.

XML provides a flexible way to create information formats. It allows structured data to be shared electronically through the Internet, Intranet, or Extranet. From this discussion, let us try to conclude the meaning of individual words in XML.

Extensible XML is extensible since it allows designers to define their own tags.

Markup The basic building blocks of XML are markups or tags.

Language XML is not just a language like HTML. Rather, it is a metalanguage—a language that allows to create or define other languages. For example, with XML, languages such as RSS, MathML (a mathematical markup language), and even tools like extensible stylesheet language transformation (XSLT) can be created.

3.7 HACKING AND ETHICAL HACKING

We are all aware of the term hacking, which means finding possible entry points in a computer system or a computer network to break into it. Hacking is an unethical and illegal activity which refers to gaining unauthorized access to a computer system or its network either to harm it or to steal sensitive information available on the computer.

What can be Hacked?

Hackers can hack the following entities:

- A website by taking unauthorized control over a web server.
- A network by gathering information about it to harm the network system and hamper its operation.
- An email account to gain unauthorized access to read all the emails or to send emails to others without taking the consent of its owner.
- A password to get unauthorized access to a computer system.
- Hackers may perform the act of hacking just for fun, show-off, to steal important information, to cause damage to the computer system, to hamper privacy, or for money extortion.



Hacking leads to massive security breach, unauthorized system access, privacy violation, hampering system operation, and malicious attack on the system.

In contrast to hacking, there is another term *ethical hacking*, which is considered legal as long as it is being done to find weaknesses in a computer or network system for testing purpose.

A computer expert who does hacking is called a *hacker*. Correspondingly, those who perform ethical hacking are known as ethical hackers. Hackers need to have knowledge to understand how systems operate and how they are designed before they play with these systems.

What is Ethical Hacking?

Till now, you have a perception that hackers are the bad guys. They are cyber criminals who steal data, blackmail victims, and use someone's confidential information for their own personal and financial gain. However, hacking is not always wrong. In fact, ethical hacking is an important method of ensuring the security of a computer system and its network.

Ethical hackers also known as white hat hackers, use the same tools and techniques that the criminal hackers (or black hat hackers) use. However, the efforts of ethical hackers are aimed at strengthening, rather than exploiting the systems they hack. They use penetration testing to find vulnerable entry points in computer network systems before the attackers can discover them. Besides discovering weaknesses, ethical hackers suggest and implement solutions to fine-tune the system's overall information security.



Vulnerabilities are found in poor or improper system configuration, hardware or software flaws, and operational weaknesses in a process.

Of course, the ethical hackers work with the permission of the organization which they are trying to defend. In many cases, the organization does not even reveal to their information security team that they have hired ethical hackers to test the effectiveness of the measures that have been implemented by them. This is referred to as a double-blind environment.

Advantages of Ethical Hacking

While hacking can be really detrimental, ethical hacking on the other hand is quite useful in certain scenarios such as the following:

- To recover lost information (like passwords)
- For doing penetration testing to find out any holes in computer and network security
- To deploy sound measures to prevent security breaches
- To use measures that prevent malicious hackers from gaining access to the computer system

Moreover, the US state government and many big companies such as IBM have especially hired a team of ethical hackers to keep their systems secure. Similarly, Trustwave Holdings Inc., has an ethical hacking lab to exploit vulnerabilities that may be present in ATMs, point-of-sale devices, and surveillance systems. Companies and governments have been using ethical hacking for the following reasons.

Fight against terrorism and national security breaches Many terrorist organizations use computer and network systems to break into various government defence systems. They collect information from these systems and then use this information to plan their activities. In such a critical scenario, ethical hackers can discover pit holes in security system and rectify them. They can also place misleading information for the attackers in the places which can be easily exploited by the attackers to divert them.

Save money by preventing loss of financial data or other confidential data The governments are saving billions of dollars by protecting their systems from the attackers. In case of a successful attack, they will have to not only compensate for the loss but also build the system from scratch which will in turn cost a lot of money. The ethical hackers test the existing defence systems to ensure a foolproof system that prevents the breakdown of the existing system.



A successful test does not mean that a network or computer system is 100% secure. But it definitely means that the system is robust against automated attacks and unskilled hackers.

Drawback of Ethical Hacking

Everything in this world has a darker side as well and ethical hacking is not an exception. Some drawbacks of ethical hacking are as follows:

- The ethical hackers must be provided complete details of the assets that should be protected and potential threat sources so that they can work effectively. What if the ethical hackers themselves misuse this information?
- Hiring an ethical hacker and trusting him/her may lead to a massive security breach.
- Even the most innocent ethical hacker may change his intentions on seeing the company's or its customers' financial and banking details.
- The ethical hacker may place malicious code, viruses, malware, and other destructive and harmful software on a computer system.

3.8 MOBILE COMPUTING

We may not be aware of the term mobile computing as such, but the truth is that now this technology has become a crucial part of our lives. *Mobile computing* is a technology that has enabled transmission of data, voice, and video using computer or any other wireless device. It involves the use of mobile communication, mobile hardware, and mobile software.

Mobile communication includes the infrastructure used to allow communication. Since the devices are wireless (with no physical links), mobile communication uses radio wave signals that are carried over the air to intended devices that are capable of receiving as well as sending similar kinds of signals.

Mobile devices use wireless network or Wi-Fi which is often referred to as a hotspot. These days, Wi-Fi can be found in offices, restaurants, educational institutions, public areas, and also at home. For wireless network, you need a wireless router connected to broadband Internet, which is often obtained through ISPs.

You must have been using wireless connection, did you notice that whenever you leave the hotspot area, your mobile device gets disconnected from the Internet? The range of Wi-Fi is usually 100 feet (it may vary with the network).

Whenever you log into Wi-Fi, the mobile device remembers passwords and automatically connects when you are in that area. For example, if you have Wi-Fi at home, then the moment you enter your home with your mobile device, it will automatically connect with the Internet.

Mobile hardware includes mobile devices such as laptops, smartphones, tablet PCs, personal digital assistants, and even the smart watches. These devices are highly portable so that you can carry them with you and access information at any time as per your convenience.

Mobile software is the actual program that runs on the mobile hardware. It is especially designed to cater to the needs of mobile applications. Some people refer to it as the engine of the mobile device. Technically speaking, mobile software is the operating system of the mobile device.

Advantages of Mobile Computing

Mobile computing is an amazing technology that allows users to stay connected to all sources at all times. Users have the power to tailor their mobile computing interface to suit their individual needs. In addition to this, other advantages of mobile computing include:

Location flexibility Since portability is the main advantage of mobile computing, the users of mobile devices do not have to be tied or pinned to a single physical location. They can operate from anywhere.

Saves time and increases productivity Mobile computing has helped users to save time, which was earlier wasted while travelling from different locations or to the office and back. Users can now access all their documents and files over the wireless connection and work as if they were on their own computer. This enhances their productivity as the users can employ the same time to do other useful work.

Ease of research Research has been made easier. Earlier users were required to do a lot of field work to search for facts, conduct surveys and then feed the data into the system. Now researchers can collect and enter data from anywhere without making unnecessary trips to and from the office to the field.

Entertainment Users can watch video and audio recordings anytime and anywhere, using mobile computing. They can see a wide variety of movies, news, sports, educational and informative material, etc., while travelling. They can even record audios and videos and store them for as long as they wish.

Streamlining of business processes Another excellent feature of mobile computing is that it ensures security through authorization and authentication. This feature helps businesses to carry out their processes over secure links and also share information between their partners.



Mobile devices also allow users to stay connected with friends and families at all times through social networking websites.

Meetings and seminars are now conducted using video and voice conferencing. This not only saves travel time but also the expenditure otherwise incurred.

Challenges and Measures

Security has always been a big concern especially in case of wireless communication. Due to its nomadic nature, it is not easy to monitor the proper usage of mobile devices. Some users with bad intentions may misuse their privilege and indulge themselves in unethical practices such as hacking, industrial espionage, pirating, online fraud, and installing malicious software to cause harm to people or devices.

Another big challenge in mobile computing is identity verification. Users just need to provide their username and password, and most of the users prefer to stay signed-in on their mobile devices. Just imagine if their devices are stolen, then all their files and data can be viewed and misused in any way by anyone.

No company would like their documents to get in the hands of hackers and other intruders, who may either misuse them or sell them to their competitors. So companies need to take certain precautions to minimize these threats. This includes hiring qualified people, installing hardware and software to monitor security, educating users on mobile computing ethics, and enforcing proper access, rights, and permissions.

3.9 INTERNET OF THINGS

Although the concept was not named until 1999, the Internet of Things has been in development for decades. *Internet of Things* (IoT) means a system of interrelated computing devices, machines, objects, animals, or people that have unique identifiers for identification and also the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction.

If the thing, in the IoT, is a person, then he can have a heart monitor implant; if it is an animal, it can have a biochip transponder; if it is an automobile, it can have sensors to generate alert alarms (for example, when the pressure in the tyre is too low); and if it is any other natural or man-made object, then it is assigned an IP address and provided with the ability to transfer data over a network.

The huge availability of address space in IPv6 has led to the development of the IoT. With more IP addresses available, more number of devices can be connected to exchange information with each other.

Today, IoT is being widely used for precision agriculture, building management, healthcare, energy and transportation. The first internet appliance was a Coke machine at Carnegie Mellon University in the early 1980s when the programmers succeeded in connecting the machine over the Internet to check the status of the machine and determine the availability of a cold drink. This check helped them to save their time as they did not have to go to the machine to buy Coke when it was not available in the machine.



In simple terminology, IoT is the concept of connecting any device over the Internet.

Later, other devices such as cellphones, coffee makers, washing machines, headphones, lamps, wearable devices, and almost anything you can think of could be connected. It is interesting to know that the analyst firm Gartner has said that by 2020, there will be over 26 billion connected devices. Other analysts have said that this number will be over 100 billion. So we can conclude that IoT is a giant network of connected “things” (including people), which will have relationships between people-people, people-things, and things-things.

Examples of Applications of IoT

Whenever we read about a new technology, the first question that comes to our mind is how this technology is going to help us? While with IoT, anything that can be connected, will be connected, but the question is why should we want these devices to be connected? There are many examples to justify the answers to these questions.

Case 1: Imagine you are going for a meeting and your car tells you the best route to take. If there is a traffic jam, then your car sends a text message to the other party notifying them that you will be late.

Case 2: Imagine that the moment your alarm rings to wake you, a message is sent to the coffee machine to start brewing coffee for you.

Case 3: The smart watch that your wear in office tells you when and where you were most active and productive.

Case 4: You have a self-driving car with complex sensors to detect objects in their path.

Case 5: A smart football that can track how far and fast it is thrown and records those statistics via an app for future training purposes.

Case 6: You have a smart refrigerator that can text message you that milk packets are over in the fridge and you need to buy them before you come home. Or, your fridge checking the expiry date of an ice-cream and notifying you that you should not eat it now.

Case 7: Imagine that you have an AC installed in your home that is connected with your smartphone. On a very hot day, you can instruct your AC to start and set the room temperature to 18 degrees before you finally reach the house.

In an article, Kevin Ashton (an innovator and consumer sensor expert) wrote that if computers knew everything about things (using data gathered by them) without any help from us then we would be able to track and count everything. It would also reduce waste, loss, and cost. We would know exactly when things need to be replaced, repaired, or recalled, and whether they were fresh or past their best.

In this way, IoT can help organizations to save a lot of money through improved process efficiency, asset utilization, and enhanced productivity. With improved tracking of objects using sensors and connectivity, companies can better analyse them and make smart decisions. For example, if you own a car manufacturing company, then you can know which accessories are the most popular by using sensors to detect which areas in the showroom are the most popular, and where customers linger the longest. You can even monitor the available sales data to identify which components are selling the fastest and then automatically align sales data with supply, so that popular items never go out of stock.



The information collected by IoT devices can be used to detect patterns, make recommendations, and detect possible problems before they occur.

We all have heard that the Indian government is working hard to develop smart cities. Do you know that smart cities use IoT for efficient utilization of resources? IoT has the potential to transform entire cities by solving real problems faced by people every day. With proper connections and data, the IoT can solve traffic congestion issues and reduce noise, crime, and pollution.

IoT Products

Big companies such as Honeywell, Hitachi, GE, Cisco, AT&T, Apple, Google, IBM, Microsoft, Skyworks, Iridium Communications, Red Hat, Zebra Technologies, and InterDigital are already playing in the market to realize the benefits of IoT. The following are some examples:

Amazon Echo for Smart Home works through its voice assistant, Alexa. Users can talk to Alexa and give order to perform a variety of functions. For example, users can tell Alexa to play music, provide a weather report, get sports scores, order an Uber, and do much more.

Fitbit One-Wearables tracks your steps, floors climbed, calories burned, and quality of sleep. The device wirelessly connects with computers and smartphones to transmit your fitness data in understandable charts to monitor your progress.

Barcelona-Smart Cities Barcelona in Spain is one of the foremost smart cities in the world. It has implemented several IoT initiatives that have helped to enhance smart parking and the environment.

AT&T-Connected Car AT&T added 1.3 million cars to its network in the second quarter of 2016. With this, the count of total number of connected cars rose to 9.5 million.



Amazon Web Services, Microsoft Azure, IBM's Watson, Cisco IoT Cloud Connect, Salesforce IoT Cloud, Oracle Integrated Cloud, and GE Predix are popular IoT platforms.

Thus, we see that IoT allows for virtually endless opportunities and challenges thereby making it a hot topic for research. According to a report, nearly \$6 trillion will be spent on IoT solutions over the next five years.

Challenges

Security is the biggest issue that we often come across while talking about IoT. With billions of devices connected together, concerns of privacy and data sharing always haunt our minds.

In fact, protection of sensitive data was ranked as the top concern among enterprises according to the 2016 Vormetric Data Threat Report. Hackers try to penetrate connected cars, critical infrastructure, and even people's homes. Therefore, the main focus of the companies is to ensure security of all the data generated by these devices.

Another issue with IoT is that massive amounts of data is being generated by these devices. So companies need to figure out how they would store, monitor, analyse, and deduce results from this vast amount of data that is continuously being generated.

However despite all odds, IoT is being used by several industries including manufacturing, defence, transportation, banks, retail, oil & gas mining, and healthcare, to name a few.

3.10 ETHICAL ISSUES IN SOCIAL NETWORKING

We are all aware of social networking sites which allow users to make their profile, find people with common interests, share information (pictures, audio, video, messages, stories, links, etc.) with friends, interact with people and lots more. The fast emergence and vast acceptance of social networking sites (SNS) by users irrespective of their age, sex or location has changed our lives completely.

SNS such as Facebook, Google+, Foursquare, and Twitter have enhanced communication capabilities by allowing users to interact with their families and friends, build business appearances or meet with other internet users. Users use these sites to share news, interest, opinions, insights, and experiences within their groups of communities. According to latest statistics released by International Telecommunication Union (ITU), almost 40% of the world's population is online and billions of Internet users use SNS. However, due to lack of awareness regarding the danger of social networking, active users are more prone to be publically targeted since they are available and searchable through their profile pages, which contain personal information such as location and phone number.

Privacy SNS has lots of data about its users. Therefore, privacy of data is a serious concern. Privacy can be violated through intrusion (attacking the network or the computer), misuse of information, and interception of information.

With the advent of Global Positioning Systems (GPS) or location-based social networking applications, users are subject to accept a data policy that might covertly take advantage of their privacy to legitimately and specifically target them with advertising ads. SNS like Facebook knows a lot of details about its users such as where they live, where they work, etc.

Do you know that many people are deleting their accounts on Facebook just for protecting their data against misuse and social pressure to add friends. This in itself is enough to demonstrate that privacy issues in SNS is in a serious and critical stage where users are unable to tolerate more and believe it is best to be totally offline.

Free speech Everyone should have freedom of speech as it is considered to be a basic human right. Though free speech is part of democratic ideology, it is a dangerous weapon for some people who misuse this right to spark unethical sentiments such as racial or religious hatred through social media.

Data leakage Data security and protection has always been the most difficult task. Although market is flooded with many new technologies that ensure data security and protection, data leakage is still a big concern. In SNS, users frequently update their current interests and location which are shown publicly. All this data can be compiled as a valuable source of information in data leakage. Even though every SNS clearly mentions their privacy policies, it depends on users how they control what information should be revealed about them and what should remain confidential.

Do you know that in 2013, Facebook disclosed that they had inadvertently exposed 6 million users' phone numbers and email addresses to unauthorized viewers over the past year and blamed the data leak on a technical problem in its archive of contact information. During this data leakage issue, Facebook users who intended to download data list of their friend's contact list would have obtained other confidential information which was never shared publicly on the network.



Users use LinkedIn to build social relations organized around professional lives, Twitter is useful for creating lines of communication between ordinary individuals and figures of public interest, MySpace was used by musicians to promote themselves and communicate with their fans. Facebook is used to connect with family, friends, and business customers.

Identity theft Identity theft is another ethical issue in SNS. In this technique, criminals impersonate someone else's identity to commit crimes so that the other person is held responsible for the consequences of fraud or crime done by them. On SNS, identity thefts are being done either by creating a fake account or stealing and hacking the password of another SNS user.

Although SNS is tirelessly working to secure and protect its users from any form of online impersonation, the best solution is always that users should have ethical values to not indulge in such activities.

Code of conduct SNS is also being used by companies to constantly interact with their current and potential customers due to which various ethical challenges have also emerged. Ethics is a very serious matter and any unethical act can lead to loss of sales and customers as well as expensive lawsuits.

Providing incorrect information, making an inappropriate comment or using SNS in an irresponsible manner can inversely effect a company's credibility. For example, if an employee replies rudely to a customer's complaint on SNS, then he is not just violating the company's code of conduct, but it will also result in negative responses from other consumers to an extent that they may even boycott the company's products. Therefore, all employees must be well trained to act with integrity and respect the company's values and code of ethics.

Respecting copyrights These days, many individuals and companies are in a mad race to write, promote, and share ideas and interesting facts on the SNS. As a result, they plagiarize and steal information from others with or without intention. For example, a company that wants to attract viewers on SNS by being the first one to release breaking news may with or without intention use pictures and footage from another source without seeking proper permissions from it or giving any credit for ownership on SNS.

Violating copyrights may result in overwhelming court cases. So, one should always provide the sources of their information when posting on SNS.

Some More Business Ethics

Unreported endorsements Giving any form of compensation to a blogger or tweeter, is considered a compensated endorsement and must be avoided. The compensation may include free product or cash. This applies to the entire social media and not just bloggers.

Compromising consumer privacy The data collected from customers is a valuable resource that companies analyse to make better decisions. But selling that data to someone else or using it for any purpose other than analysing or providing better service to the customers is unethical and an act of violating consumer's privacy. Even when these breaches are unintentional and non-malicious, they scare consumers away.

Affecting company's brand value Today, individuals are on SNS even before they join any company. Prior to joining the company, they may have had their own opinions but once they join an organization, they must remember two things. First, if they are promoting products of their company, they should also specify their connection to the company as otherwise they would be violating consumer's trust. Second, they should never criticize their company and never disclose their company's policies on SNS.

Using the online community to get free work These days, while companies use the term UGC (user generated content), but in creative circles it is being known as WFF (working for free). Basically, more and more companies are holding online contests allowing its customers to submit designs or even produced commercials. They then pick a winner, giving him/her prizes and using his/her creative work in traditional media. However, the truth is that they are using the skills of freelancers or even unemployed professionals who should be paid for their work.

Another trend is that companies request for proposals for a social media campaign, gather the most creative ideas that come out of the process and use them without hiring the particular agencies who created them.

Giving false information For social media marketing, it is the ethical responsibility of the company to be honest and truthful regarding posting product information and pricing.

Never disclose consumer's information It is unethical to reveal any information about consumers on SNS. A company should be considerate of its clients' privacy issues and any kind of personal information and purchasing details must never be revealed in public.

Cybercrime on SNS Many types of cybercrimes such as cyberbullying/cyber harassment, cyber stalking, child exploitation, cyber extortion, cyber fraud, illegal surveillance, cyber espionage, cyber sabotage, and cyber terrorism are reported on SNS. Such activities are not just unethical but also anti-social.

SNS is caught between the public interest in crime prevention and their need to preserve the trust and loyalty of their users as users feel that government is violating their privacy and keeping a check on their online activity.

Many companies prefer user security by employing end-to-end encryption of messages exchanged on SNS to prevent government agencies from accessing user data in the interest of public safety and national security.

Overwhelming promotional messages Repeatedly displaying unmasked promotional messages (advertisements) is also an unethical act as it annoys the users.

Dishonesty and distortion It is unethical to be dishonest about anything even on social networks. So, making false claims about yourself or anything else or posting offensive material has an adverse effect on one's reputation. So, one should talk on SNS keeping ethics in mind.

Improper anonymity and distorted endorsements Giving wrong affiliations, credentials or expertise, is unethical. Similarly, giving anonymous and fake feedbacks just to cause damage to a company's reputation is also considered to be unethical.

Many companies even hire people to comment on fabricated stories about their products is another example where ethics go downhill.

In fact, due to the rise of mobile social networking, the concern for ethical values has become even more severe as everyone is connected to his/her family members, neighbours, friends, relatives, and colleagues on SNS. So any unethical act on SNS can affect one in a big way, both in personal as well as professional life.

Summary

- The Internet is a global network that connects billions of computers all over the world.
- Electronic mail (email) means the transmission of messages over communication networks.
- Internet chat allows two or more online users to come together to talk using an instant messenger.
- With online newspaper, users can read the full coverage of breaking news in a timely manner.
- Online shopping means buying goods and/or services from merchants who sell on the Internet.
- The contents hosted on the Intranet can be accessed only by members within the organization who have appropriate access control rights.
- Internet protocol address is a unique address allotted to computing devices such as computers, routers, printers, scanners, modems, smartphones, tablets, and so on that are connected to the Internet.
- The term 'protocol' means a set of rules that must be followed to facilitate communication.
- Domain name system (DNS) is a service that automatically converts domain names into IP addresses.
- Transmission control protocol (TCP) is a protocol that works at the transport layer and is used with the IP protocol which is used to send data packets between sender and receiver devices.
- While an absolute uniform resource locator (URL) specifies the complete URL containing all three fields—protocol, domain, and path, relative URLs, on the other hand, contain only one field—the domain name.
- Dynamic HTML is a combination of technologies—HTML, Java Script, and Cascading Style Sheets and is used to create dynamic, interactive, and animated web pages.
- Extensible markup language (XML) is a markup language that defines a set of rules for encoding documents in a format that is readable by humans as well as computers. It is used to share information in a consistent way.

Glossary

Browser Software that enables users to access the World Wide Web

Chat An Internet service that enables users to communicate with other people in real time. The user

types his message using the keyboard. The message then appears on the screen of the other person. Similarly, the message typed by the other person appears on the user's screen

Client A program that requests information from another computer on the network. The client program accesses and displays the retrieved information to the user. For example, the web browser is a client program that accesses and displays web pages

Download Making a copy of a file from another computer to your computer. For example, users may copy a song from another computer on the Internet to his computer

Electronic mail An email system designed to enable the users to send and receive messages across a network

FTP Short form of file transfer protocol, it is basically a program that enables users to transfer files from one computer to another

Gopher A program that organizes information on the Internet using a system of menus. Items in the menus can be links to other documents or to other information services.

Host Any computer on a network

HTML HTML stands for hypertext markup language and is used to create hypertext documents for use on the World Wide Web

HTTP Hypertext transfer protocol (HTTP) is a protocol that defines a set of rules to exchange documents on the World Wide Web

Hypertext A hypertext document is one that includes links to other documents on the World Wide Web

Internet A network of networks. It connects several networks all around the world to enable them to exchange information with each other. For this purpose, all the computers on the Internet use a common set of rules (protocols) for communication. Therefore, the Internet uses a set of protocols called transmission control protocol/Internet protocol

Internet service provider A commercial service that sells access to the Internet to individuals. Users connect to the ISP through a modem. While some ISPs only offer a basic connection to the Internet, others, on the other hand, sell a variety of value-added services such as discussion forums, tech

support, software libraries, news, weather reports, stock prices, plane reservations, and even electronic shopping malls

Link A word, picture, or other area of a web page that users can click on to move to another spot in the same document or to another document. Links (words) may be underlined and usually appear in a contrasting coloured text. When the user clicks on the link, the colour of the text changes

Network A group of computers connected to each other to exchange information and/or resources

Newsgroup An Internet service in which readers can post messages or articles for other people to read. Other people can also reply to articles that they read on a newsgroup. It enables people with similar interests to communicate with each other

Node Computer connected to a network

Protocol A set of standardized rules that should be followed to exchange information among computers. There are different protocols for different kinds of communication. For example, HTTP specifies the rules for exchanging information on the World Wide Web. FTP defines the rules to copy files from one computer to another across a network

Server A program that provides information or services to other programs. For example, the web browser is a client that uses services like email from the server

Upload The opposite of download that means transferring a file, picture, document, or an audio/video clip from your computer to another computer

URL A uniform resource locator specifies the addresses for World Wide Web pages. A URL uniquely identifies a Web page. URLs have three parts—protocol name, server name, and a directory path. For example, consider the URL <http://wings.avkids.com/SPIT/index.html>. Here, <http://> is the name of the protocol, wings.avkids.com is the server's name, and [/SPIT/index.html](http://wings.avkids.com/SPIT/index.html) is the location of the file on the server

Web page A document on the World Wide Web that can contain text, pictures, movies, sounds, or links to other pages

Website A collection of web pages on the World Wide Web having to do with a particular topic or organization

World Wide Web An interconnected set of hypertext documents located throughout the Internet

Multiple-choice Questions

1. _____ is a global network that connects billions of computers all over the world.
 - (a) LAN
 - (b) MAN
 - (c) WAN
 - (d) Internet
2. _____ is a network of networks.
 - (a) Internet
 - (b) Intranet
 - (c) Extranet
 - (d) All of these
3. Each computer on the Internet is called a _____.
 - (a) server
 - (b) web page
 - (c) device
 - (d) host
4. _____ provides internet service.
 - (a) ASP
 - (b) ISP
 - (c) JSP
 - (d) IPS
5. ISP stands for _____.
 - (a) internet service provider
 - (b) internet service program
 - (c) intranet service provider
 - (d) intranet service program
6. You can share _____ using the Internet.
 - (a) voice messages
 - (b) text messages
 - (c) video messages
 - (d) multimedia messages
7. Mozilla Firefox and Opera are examples of _____.
 - (a) ISPs
 - (b) web browsers
 - (c) email applications
 - (d) social networking sites
8. _____ allow users to browse the Internet.
 - (a) ISPs
 - (b) Web browsers
 - (c) Email applications
 - (d) Social networking sites
9. The first ever network of computers was created for _____.
 - (a) military purposes
 - (b) business purposes
 - (c) research organizations and universities
 - (d) all of these
10. The first network of computers was known as _____.
 - (a) ARPANET
 - (b) APARNET
 - (c) PICONET
 - (d) Internet
11. _____ means transmission of messages electronically over communication networks.
 - (a) Courier services
 - (b) RTGS
 - (c) Email
 - (d) SMS
12. The address of the email recipient is written in the _____ field.
 - (a) From
 - (b) To
 - (c) Subject
 - (d) Body
13. Content of the message is written in the _____ field of email.
 - (a) From
 - (b) To
 - (c) Subject
 - (d) Body
14. The two optional fields of a message are _____.
 - (a) To and From
 - (b) CC and BCC
 - (c) Subject and Body
 - (d) None of these
15. When you have to send the same message to multiple recipients, then you should ideally mention all the email addresses (except one) in the _____ field.
 - (a) To
 - (b) CC
 - (c) BCC
 - (d) From

16. The _____ option allows users to make a copy of a message received from a person and then email it to someone else.
 (a) Send (c) Forward
 (b) Reply (d) Post
17. The _____ application allows users to exchange files.
 (a) SMTP (c) UDP
 (b) FTP (d) IP
18. _____ is the application using which two users can talk instantly.
 (a) Email (c) Browsing
 (b) Chatting (d) Transferring
19. Which of the following is not a reason for businesses to use instant chatting?
 (a) Answer user queries
 (b) Provide online support
 (c) Conduct business meetings
 (d) None of these
20. During Internet conferencing users cannot share _____.
 (a) information (c) video
 (b) files (d) none of these
21. _____ is not required for Internet conferencing.
 (a) Webcam
 (b) Microphone
 (c) Chatting software
 (d) None of these
22. Newspaper on the Internet is also known as _____.
 (a) online newspaper
 (b) electronic newspaper
 (c) web newspaper
 (d) all of these
23. HTML is a _____.
 (a) language (c) application
 (b) protocol (d) website
24. _____ allows users to add links to other documents, graphics, audio, and/or video files.
 (a) FTP (c) HTML
 (b) Telnet (d) HTTP
25. The text on which the mouse pointer changes to hand-shaped is known as _____.
 (a) high text (c) link text
 (b) hypertext (d) anchor text
26. WWW was created by _____.
 (a) Larry Page
 (b) Bill Gates
 (c) Tim Berners-Lee
 (d) Steve Jobs
27. Which of the following is not correct about online shopping?
 (a) You can make the payment through debit/credit card.
 (b) It can be done 24×7.
 (c) You can touch and feel the products sold online.
 (d) There is always a chance of your confidential information being compromised.
28. IP address is allotted to _____.
 (a) computers (c) printers
 (b) smartphones (d) all of these
29. _____ facilitates unique identification of devices for communication over the Internet.
 (a) IP address (c) MAC address
 (b) Sequence number (d) Port number
30. The IP address has _____ parts.
 (a) 1 (c) 3
 (b) 2 (d) 4
31. Each part of the IP address has a number ranging from _____.
 (a) 0-999 (c) 0-255
 (b) 0-128 (d) 0-256
32. _____ address is issued using a leasing system.
 (a) Static IP (c) Fixed IP
 (b) Dynamic IP (d) Permanent IP
33. _____ address reveals technical information about the continent, country, and city in which the computer is located.
 (a) Static IP (c) Temporary IP
 (b) Dynamic IP (d) Momentary IP
34. There are limited _____ IP addresses.
 (a) static (c) temporary
 (b) dynamic (d) momentary
35. Web servers are allotted _____ IP address.

- (a) static (c) temporary
(b) dynamic (d) momentary
36. _____ service translates domain names into IP addresses.
(a) TCP (c) DNS
(b) IP (d) UDP
37. DNS stands for _____.
(a) domain name system
(b) data name server
(c) data number system
(d) domain number system
38. TCP stands for _____.
(a) transmission carrier protocol
(b) transmission control protocol
(c) transmission control program
(d) transmission carrier program
39. Which of the following statements is not true about DNS?
(a) IP addresses should be assigned statically.
(b) Name and address pair can be cached.
(c) ISPs have their own DNS server.
(d) None of these
40. The domain .in belongs to which country?
(a) Indonesia (c) Iceland
(b) India (d) Iran
41. Which of the following is true about domain names?
(a) org is used for non-profit organizations.
(b) Domain names are also organized from right to left.
(c) mobil is a valid domain name for mobile communication network
(d) All of these
42. _____ specifies the unique address for a file that is accessible on the Internet.
(a) URL (c) UDP/IP
(b) TCP/IP (d) DNS
43. The syntax for specifying a URL is _____.
(a) protocol://domain-name/path
(b) path://domain-name/protocol
(c) domain-name://protocol/path
(d) protocol/domain-name/path
44. Relative URL contains only the _____ field.
(a) protocol (c) domain name
(b) path (d) none of these
45. Anything following the question mark (?) in a URL is a _____.
(a) domain-name
(b) protocol
(c) variable-value pair
(d) path
46. In a URL, two variable-value pairs are separated with a/an _____ symbol.
(a) ! (c) @
(b) & (d) +
47. The current version of IP is _____.
(a) 4 (c) 6
(b) 5 (d) 7
48. IPv6 has a _____ bit address.
(a) 32 (c) 128
(b) 64 (d) 256
49. _____ protocol establishes a connection between the sender and the receiver so that reliable and error-free data transmission can take place.
(a) TCP (c) UDP
(b) IP (d) SMTP
50. Segmentation and reassembly is done by which protocol?
(a) TCP (c) UDP
(b) IP (d) SMTP
51. _____ is an alternative protocol to TCP.
(a) FTP (c) UDP
(b) IP (d) SMTP
52. Which transmission protocol is best-suited for online computer games, watching videos online, and listening to audio online?
(a) FTP (c) UDP
(b) IP (d) SMTP
53. Identify the incorrect statement about UDP.
(a) It is preferable for small data exchanges.
(b) It does nothing to correct errors and supports unreliable data transfer.
(c) It is a connection-oriented protocol.
(d) It does not break the message into smaller segments.

54. FTP is not _____.
(a) fast (c) efficient
(b) reliable (d) connection-less
55. FTP allows users to _____ files.
(a) rename (c) delete
(b) copy (d) all of these
56. In which mode does the client initiate the connection?
(a) Active connection mode
(b) Passive connection mode
(c) Block mode
(d) Stream mode
57. In which mode is the server always waiting for any request from the client?
(a) Active connection mode
(b) Passive connection mode
(c) Block mode
(d) Stream mode
58. In which mode are files transferred as a continuous stream with no intervention?
(a) Active connection mode
(b) Passive connection mode
(c) Block mode
(d) Stream mode
59. _____ is a text-based protocol that is used for accessing a remote computer's (called host) data and application programs.
(a) FTP (c) UDP
(b) Telnet (d) SMTP
60. _____ enables research scholars and professors to log in to the university's computer from any terminal.
(a) FTP (c) UDP
(b) Telnet (d) SMTP
61. Which of the following is incorrect about Telnet?
(a) It is a text-based computer protocol.
(b) It is insecure because it transfers all data in clear text.
(c) Users cannot transfer files using Telnet.
(d) None of these
62. _____ is a code embedded in a file that instructs the web browser how to display the page.
(a) Hyperlink (c) Markup
(b) Hypertext (d) Image
63. _____ interprets the meaning of markups.
(a) HTML (c) Web server
(b) Web browser (d) Web client
64. Which of the following cannot be embedded in an HTML page?
(a) JavaScript (c) Visual Basic
(b) VBScript (d) Form
65. Which of the following is best-suited for creating an interactive game?
(a) HTML (c) CSS
(b) JavaScript (d) DHTML
66. Using which language can you create your own tags?
(a) HTML (c) XML
(b) DHTML (d) All of these
67. Using which language can arbitrary data structures be easily represented?
(a) HTML (c) XML
(b) DHTML (d) All of these
68. Nested tags are used in which language to represent hierarchical data?
(a) HTML (c) XML
(b) DHTML (d) All of these
69. In which language are tags not pre-defined?
(a) HTML (c) XML
(b) DHTML (d) All of these
70. _____ means finding possible entry points in a computer system or a computer network to break into it.
(a) Hacking (c) Cracking
(b) Ethical hacking (d) Ethical cracking
71. _____ is done to steal sensitive information available on the computer.
(a) Hacking (c) Cracking
(b) Ethical hacking (d) Ethical Cracking
72. Hacking does not involve _____.
(a) breaking passwords
(b) recovering passwords
(c) illegal use of someone else's email account
(d) harming a computer system

73. _____ means finding weaknesses in a computer or network system for testing purpose.
- (a) Hacking (c) Cracking
(b) Ethical hacking (d) Ethical cracking
74. _____ are known as black hat hackers.
- (a) Ethical hackers (c) Crackers
(b) Hackers (d) Ethical crackers
75. Ethical hackers are also known as _____ hackers.
- (a) white hat (c) black hat
(b) red hat (d) grey hat
76. Vulnerabilities are present in a computer due to _____.
- (a) improper system configuration
(b) hardware or software flaws
(c) operational weaknesses in a process
(d) all of these
77. When an organization does not even reveal to their information security team that they have hired ethical hackers to test the effectiveness of the measures taken by them, then this is called _____ environment.
- (a) blind (c) double blind
(b) single blind (d) All of these
78. Which of the following statements is true about ethical hackers?
- (a) Ethical hackers must be provided complete details of the assets.
(b) Hiring an ethical hacker and trusting him may lead to a massive security breach.
(c) The ethical hacker may place malicious code, viruses, malware, and other destructive and harmful software on a computer system.
(d) All of these
79. Mobile communication uses _____ signals that are carried over the air to intended devices.
- (a) ultraviolet wave
(b) infrared wave
(c) radio wave
(d) microwave
80. The main components of mobile computing are _____.
- (a) mobile hardware
(b) mobile software
(c) mobile communication
(d) all of these
81. Which of the following is not an advantage of mobile computing?
- (a) Highly secure
(b) Enhances productivity
(c) Location flexibility
(d) Entertainment
82. IoT connects different _____.
- (a) objects (c) animals
(b) machines (d) all of these
83. _____ is the concept of connecting any device over the Internet.
- (a) Mobile computing
(b) IoT
(c) Internet
(d) Social networking sites
84. The technology used in self-driving car, smart watch, smart football, and smart refrigerator is _____.
- (a) mobile computing (c) Internet
(b) IoT (d) Android
85. Google+, Foursquare, Pinterst, and Twitter are all examples of _____.
- (a) search engines
(b) social networking sites
(c) video conferencing sites
(d) instant messaging sites
86. _____ sites are used to share news, interest, opinions, insights, and experiences.
- (a) Search engines
(b) Social networking sites
(c) Video conferencing sites
(d) Instant messaging sites
87. Challenge(s) in SNS include _____.
- (a) free speech (c) code of conduct
(b) data leakage (d) all of these
88. Which SNS is used to build social relations organized around professional lives?
- (a) LinkedIn (c) Facebook

89. _____ is used for creating lines of communication between ordinary individuals and figures of public interest.
 (a) LinkedIn (b) MySpace (c) Facebook (d) Twitter
90. If anyone misbehaves on SNS, then it means _____.
 (a) misuse of free speech (b) data leakage (c) violation of code of conduct (d) fighting on the Internet
91. Not providing the sources of their information when posting on SNS results in _____.
 (a) misuse of free speech (b) data leakage (c) violation of copyrights (d) violating code of conduct
92. Giving any form of compensation to a blogger or tweeter is considered as _____.
 (a) unreported endorsements (b) compromising privacy (c) effecting company's brand value (d) violating code of conduct
93. The term UGC refers to _____.
 (a) user generated content (b) getting creative ideas (c) holding online contests allowing its customers to submit designs (d) all of these
94. Giving anonymous and fake feedback just to cause damage to a company's reputation is known as _____.
 (a) improper anonymity (b) distorted endorsements (c) cybercrime (d) identity theft

Answers to Multiple-choice Questions

- | | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1. (d) | 2. (a) | 3. (d) | 4. (b) | 5. (a) | 6. (d) | 7. (b) | 8. (b) | 9. (c) | 10. (a) |
| 11. (c) | 12. (b) | 13. (d) | 14. (b) | 15. (c) | 16. (c) | 17. (b) | 18. (b) | 19. (d) | 20. (d) |
| 21. (d) | 22. (d) | 23. (a) | 24. (c) | 25. (b) | 26. (c) | 27. (c) | 28. (d) | 29. (a) | 30. (d) |
| 31. (c) | 32. (b) | 33. (a) | 34. (a) | 35. (a) | 36. (c) | 37. (a) | 38. (b) | 39. (d) | 40. (b) |
| 41. (d) | 42. (a) | 43. (a) | 44. (c) | 45. (c) | 46. (b) | 47. (c) | 48. (c) | 49. (a) | 50. (a) |
| 51. (c) | 52. (c) | 53. (c) | 54. (d) | 55. (d) | 56. (a) | 57. (b) | 58. (d) | 59. (b) | 60. (b) |
| 61. (d) | 62. (c) | 63. (b) | 64. (c) | 65. (d) | 66. (c) | 67. (c) | 68. (c) | 69. (c) | 70. (a) |
| 71. (a) | 72. (b) | 73. (b) | 74. (b) | 75. (a) | 76. (d) | 77. (c) | 78. (d) | 79. (c) | 80. (d) |
| 81. (a) | 82. (d) | 83. (b) | 84. (b) | 85. (b) | 86. (b) | 87. (d) | 88. (a) | 89. (d) | 90. (c) |
| 91. (d) | 92. (a) | 93. (d) | 94. (b) | | | | | | |

Security and Encryption

CHAPTER

4

Syllabus Mapping	Unit
Need and concepts, dimension, definition and scope of e-security, security threats—malicious codes (virus, Trojan horse, worm, spyware, ransomware), hacking, spoofing, sniffing, phishing, spamming, denial of service (DoS) attacks, technology solutions [confidentiality: (data encryption & decryption, symmetric and asymmetric encryption), security implementation: firewall, DMZ (de-militarized zone), SSL, HTTPs, significance of website auditing].	Module I Unit 4

4.1 NEED AND CONCEPTS

Data security is a broad issue that encompasses security for all the data and information that an organization stores on the computer and security of all the transactions that are made using the Internet. It ensures authenticated access of data. The term 'data security' is gaining topmost priority, especially in financial and government institutions as lack of security is a serious threat to the integrity and privacy of any organization.

Since the Internet is an insecure channel for exchanging private data or messages and intrusion or frauds like phishing (discussed later) are very common, some methods must be implemented to protect data. In this chapter, we will read about the threats to security and learn about the protective measures that can help the users protect their data from unwanted access. Before delving into these issues, let us try to find answers to some important questions.

How can Security of Data be Compromised?

Security of data may be compromised in the following ways:

- Unauthorized users from within or from outside the organization may access the data.
- Authorized or unauthorized users may modify the existing data, add wrong data, or delete some important data.

Figure 4.1 reveals the sources that can cause threat to security of data. While we cannot control damage to data because of natural disasters, damage due to humans can always be controlled by implementing sound security mechanisms. A hacker is someone who either breaks into the system for which he has no authorization or goes beyond his limits of

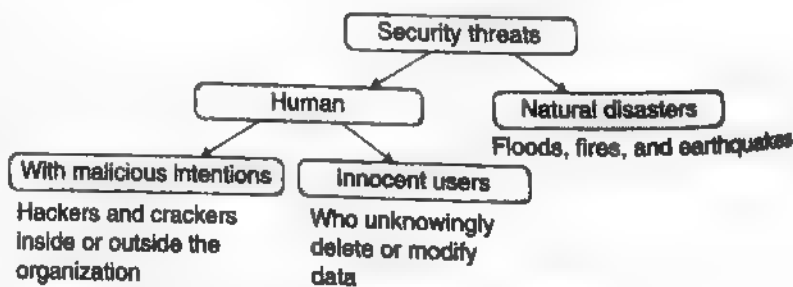


Figure 4.1 Types of security threats

legitimate access. A hacker can be a cracker; a cracker means a person who breaks into the system by password cracking or by cracking the security measures implemented to protect the data.

However, a big threat to data security can be employees within the organization who do not have any malicious intention, but they either unknowingly alter data themselves or leave the data vulnerable to be accessed by attackers. For example, a programmer can write a code that can result in system crash. Another example is a clerk who saves an important file on the desktop. This file can then be accessed by any other user with malicious intentions.

How Is Confidential Data Tampered?

Confidential data may be tampered in the following ways:

- Delete vital information from the database—for example, if someone deletes the records of your customers, there is no means to contact them.
- Steal information—for example, if your competitor steals information about your customers and offer them goods and services at a heavy discount, then you will have to bear a great loss.
- Modify data—for example, if a student gets access to the master computer of your school and modifies his marks, then the whole concept of taking exams and evaluating students will become a wasteful exercise. Another example could be that an employee can gain access to the attendance records and mark himself present on days when he had not gone to office.

How can we Protect Confidential Data?

The following are the ways to protect confidential data:

- Prevent the data from being altered or accessed by people with malicious intentions by implementing sound security policies.
- Detect any attempt to damage, modify, or steal data. Various tools are available in the market that detect and reveal any attempt made to breach data security.
- Recover lost or damaged data if someone has already tampered with it.

4.1.1 Dimension

We all agree to the point that the number of Internet users in India is constantly increasing. Computers, smartphones, and the Internet are now becoming widespread even in rural India. As a result, cyber-crimes are also on a constant rise because the criminal no longer needs to be physically present when committing a crime. Criminals take full advantage of the Internet's speed, convenience, anonymity, and lack of borders either to carry out financial crimes such as theft, money laundering, and fraud or hate crimes, such as stalking and bullying.

By the term cybercrime, we mean any crime or illegal activity that involves a computer and a network. Cybercrimes are usually divided into the following three categories:

1. Crimes in which a computer is the target. For example, to gain network access.
2. Crimes in which the computer is used as a weapon. For example, to launch another attack.
3. Crimes in which the computer is used as an accessory to a crime. For example, using a computer to store illegally-obtained data.

Cybercrime may be carried out by individuals or by a group of people who usually target people in countries with weak or non-existent cybercrime laws.

Cybercrime also includes identity theft, threatening a nation's security, copyright infringement, and child pornography. Therefore, these crimes have become a threat to an individual's privacy. In some cases, government organizations have also been found involved in cybercrime. For example, the National Security Agency (NSA) of the United States was held responsible for intercepting and spying on millions of online users.

Moreover, in 2013, Chinese hackers were caught hacking into big US organizations and agencies like the attack on the New York Times. Cybercrimes which involve nations and state machinery are often termed as cyberwarfare.

4.2 DEFINITION AND SCOPE OF E-SECURITY

The term e-security includes a set of technologies, processes, and practices designed to protect networks, computers, programs, and data from any kind of cyber-attack, damage, or unauthorized access.

These days, e-security has become even more difficult because cyber criminals have learned to launch automated and sophisticated attacks at lower costs. Before moving into details of security, let us first try to understand some important terms.

Attack or security breach An attempt to gain unauthorized access or deny authorized access to a resource.

Availability The degree to which data and other resources are accessible for use when required.

Compromise An unauthorized access to steal, modify, or destroy confidential data.

Confidentiality The degree to which confidential data is protected from unauthorized access.

Integrity The degree to which the accuracy and completeness of information and computer software are protected against unauthorized access.

Authentication Confirming the identity of the person requesting to gain access to a resource. This can be done by validating username and password.

Authorization Determining the set of actions that an authenticated person can perform on a resource. For example, an employee working in the Accounts department is not authorized to access the files of the Sales & Marketing department.

Accountability The activities of the authorized persons are documented to deter employees from wrong doing.

Vulnerability assessment An audit or a check that is performed to identify potential vulnerabilities in a computer system or network.

Cyberterrorism Disruptive use of IT by terrorists to attack on networks, computer systems, and telecommunication infrastructures.

Cyberwarfare Nations using IT to penetrate another nation's networks to cause damage or disruption. Cyber warfare has been ranked as the fifth domain of warfare (following land, sea, air, and space). It is done to steal confidential data, degrade communications, hamper infrastructural services such as transportation and medical services, or other commercial activities.

Cyberespionage The practice of using IT to obtain secret information without seeking permission from its owners. It is usually done using cracking techniques and malware.

Cyberstalking It is the use of the Internet to stalk or harass an individual, an organization, or a specific group. Cyberstalking also includes monitoring someone's online activity.

4.3 THREATS TO DATA SECURITY

The most common threats to data security come either from use of malwares or through fraud like phishing. The common threats to security can be classified as follows.

4.3.1 Malware

Malware or malicious software (meaning software designed with wrong intentions) is software specifically designed to gain access to a computer either to disrupt its operation or gather sensitive data from it. It is a big threat to Internet security and includes computer virus, spyware, worms, Trojan horse, etc. Malwares are usually embedded within legitimate software that is either useful or attractive.

Virus

A computer virus is a small program that gets loaded in the computer without the user's knowledge and replicates itself repeatedly. Such a piece of code is dangerous because it will quickly use all available memory and finally halt the system. An even more dangerous type of virus may corrupt or delete files from the computer and may spread itself to other computers by using the user's email program. Sources of computer virus include the following:

- Attachments in email messages or instant messaging messages
- Attachments of funny images, greeting cards, or audio and video files
- Downloads from the Internet

Computer viruses are thus always undesirable as they slow down the computer's performance, cause erratic behaviour, loss of data, and frequent crashes.

Features The features of virus are as follows:

- It replicates itself.
- It requires a host file to spread.
- It is activated by an external action.

Types of virus Viruses are of the following types.

Boot viruses These viruses were used to infect floppy disks. However, as floppy disks are no longer used, the boot virus infects only the master boot records of the hard disk. The boot record is a program that loads the operating system in the memory when the computer is turned on.

Boot record viruses either overwrite or replace the boot record and move it to a different location in the hard disk. When the operating system is loaded in memory, the virus also gets loaded along with the operating system.

Once the virus gets loaded in the memory, it performs its intended task. The only way to protect the computer in case of boot virus is to boot the operating system from another hard drive or a bootable CD or DVD. Examples of boot viruses include Polyboot.B, AntiEXE, Disk Killer, and Stoned.



The master boot record is the first sector on a partitioned storage device.

Program viruses or file infector viruses These viruses infect only executable files with extensions such as .BIN, .COM, .EXE, .OVL, .DRV, and .SYS. When an infected file is executed, the file along with the virus gets loaded in the memory. The virus is then free to perform its intended task. The virus may overwrite the infected file or replace some parts of the file in such a way that every time the program

or file is intended to execute, the virus gets executed. Most of the viruses belong to this category. However, program viruses are better than boot viruses as they can be removed easily, for example, Snow.A, Jerusalem, and Cascade.



The program or file virus uses a file as its host as it sticks to a file.

Multipartite viruses These viruses are a combination of two types of viruses—boot viruses and program viruses. Like program viruses, they infect an executable file and when the infected file is executed, the multipartite virus infects the master boot records (as the boot virus). For example, Emperor, Anthrax, Tequilla, and One_Half.

Stealth viruses These viruses use several techniques to avoid their detection. For example, stealth virus will remove the virus code from an infected file when antivirus software is examining the system so that the file is not detected as infected.

Some stealth viruses redirect the hard disk head so that the next read operation is done from another memory sector instead of the correct one. Some may alter the attributes of the infected file.

Polymorphic viruses These viruses create copies during replication. These replicated copies are functionally equivalent but have different codes. The difference in codes is intentionally induced by randomly inserting superfluous instructions, changing the order of instructions, or by choosing a different encryption algorithm each time a copy of the virus is created.

The real power of this virus lies in the fact that since each infection is different from the other, it is very difficult for an antivirus software to identify, locate, and remove them. Examples include Elkern, Marburg, Satan Bug, and Tuareg.

Macro viruses A macro virus infects documents that contain macros (a special type of program that performs a series of operations with a single action). A wide variety of programs, including Microsoft Word and Excel, support macros, and are thus, vulnerable to macro viruses.

For example, many a time, when working with Microsoft Word, you must have encountered a message as 'Problem in Normal.dot'. Macro virus infects normal.dot which is a file used by all the documents. Whenever users open any document (by indirectly using Normal.dot), the uninfected document also gets infected. When a macro virus-infected file is opened on another computer, the virus spreads on that computer too. Examples of macro viruses are Relax, Melissa.A, Bablas, O97M/Y2K, and WM.NiceDay.

ActiveX viruses Most Internet users do not know how to configure ActiveX and Java controls and thus leave a security hole in their computer. Many times, while surfing the Internet, we get a pop-up message saying 'Applets are not able to run', and asking 'allow or disallow?' Sometimes, certain websites ask users to download certain ActiveX or Java controls and we quickly download them.

By allowing these applications to run freely on our machines, we permit them to deliver all ActiveX viruses. Therefore, by simply turning off some ActiveX and Java controls in the browser, computers can be protected from such macro viruses.

Resident viruses A resident virus inserts itself in the computer's memory (RAM). From the memory itself, it performs all its intended tasks such as interrupting the system's operations and corrupting files and programs that are opened. A resident virus runs independently of the file that was originally infected. Examples include Randex, CMJ, Meve, and Mr Klunky.

Direct action viruses This type of virus replicates itself and performs its intended action only when the infected file is being executed. When the file is not being executed, the virus becomes dormant. Direct action viruses take action when a specific condition is met. When the virus becomes active, it infects all the files

in its directory as well as in the directories specified in the AUTOEXEC.BAT file PATH. An example is the Vienna virus.



AUTOEXEC.BAT is a batch file stored in the root directory of the hard disk. It is used to perform some vital operations when the computer is booted.

Overwrite viruses This virus deletes the data stored in the infected files leaving them partially or totally useless. To clean a file infected by an overwrite virus, users have no option but to delete the file completely, thereby losing all its contents. Examples are Way, Trj.Reboot, and Trivial.88.D.

Directory virus A directory virus changes the paths that indicate the location of a file. When the user executes a directory virus-infected file (having extension as .EXE or .COM), he unknowingly runs the virus since the original file has been moved to another location by the virus.

Network virus These viruses rapidly spread through a LAN or through the Internet. Network viruses multiply through shared resources such as shared drives and files. When a computer gets infected, it searches through the network to attack another computer. When the other computer gets infected, it moves on to the next, and so on. Examples are Nimda and SQL Slammer.

Space filler (Cavity) viruses We have seen in Chapter 4 that many a time, some parts of a file are empty. The space filler virus use this empty space to house (or install) its code. It does not affect or damage the contents of the actual program itself. An example is the Lehigh virus.

FAT virus File allocation table (FAT) is a table maintained by the operating system to store information about location, size, and other details of files on the disk. The FAT virus attacks on the file allocation table.

The FAT virus makes it impossible for a computer to locate files. The virus spreads to the files when the FAT attempts to access them, thereby penetrating into the entire computer. When a file gets infected, to the users, it seems as if the file is missing or inaccessible.

The FAT virus disrupts a system completely by destroying data and forcing the user to reformat the system.

Worms

Like viruses, worms are programs written with malicious intentions that can replicate themselves and spread across a computer network. However, unlike viruses, most worms do not interfere with the normal use of a computer. Moreover, they exist as separate entities and do not attach themselves with other files or programs.

However, worms may also take control over the computers on which they get installed and steal confidential data. Once a worm gets installed, it uses the email program of the user to send a copy of itself to everyone listed in his email address book. Then, it replicates itself to send itself out to everyone listed in each of the receiver's address book, and continues the replication process indefinitely. Worms use a lot of network bandwidth and memory which causes web servers, network servers, and individual computers to stop responding.



Once installed, a worm can connect to a remote computer over the Internet to download a more substantial piece of malicious software.

How a worm works A worm is a malicious code that locates vulnerabilities on a computer to exploit them. Once a computer gets infected, the worm will find other computers on its network with vulnerabilities so that they can also be accessed and infected.

Worms also spread through email attachments. To users, it seems as if the email has come from a known trusted person and the moment the user opens the email, the worm uses the user's email account and address book to copy itself and spread to other email recipients.

Worms can also attack applications such as Microsoft Word and Excel by inserting malicious code in documents and then use them as an attachment with email. The most destructive feature of a worm is that it can replicate itself 2,50,000 times over a period of several hours. Besides looking for other computers, worms also scan for unsecured servers and then replicate themselves on each server. Some worms are specifically designed to replicate themselves on specific days for making targeted attacks on special occasions.

The main aim of worms is to slow down the Internet due to the massive amount of traffic it creates. Worms can also gain unauthorized access to a website to attack it by sending thousands of requests in order to crash the site.



Worms also spread through pirated movies.

How to protect your computer from worms? The following are the ways to protect the computer from worms:

- Always use the latest operating system.
- Install antivirus software to remove the worm.
- Install a firewall that will guard against the downloading of the worm.
- Update your antivirus software on a regular basis.

Trojan Horse

A Trojan horse is a non-self-replicating malicious software that pretends to be harmless so that users can easily download it on the computer. It is usually contained inside a harmless program. Once executed, a Trojan may slow down the computer, cause loss or theft of data, give unauthorized access to its controller, ruin the FAT, and install a virus.

Types of Trojan horses Once a Trojan horse gets installed on your computer, its range of actions can vary from being harmless to destructive. For example, a Trojan can cause the following:

- It might display annoying messages on the screen.
- It can delete all vital files.
- It can steal confidential information like passwords.
- It may install viruses or another Trojan horse on the computer.
- It may allow the computer to be accessed from a remote machine.

The following are some common types of Trojan horses.

Remote administration Trojan horse This type of Trojan horse gives control of the infected computer to a hacker who can alter the registry, rearrange folders, change the login password, upload or download files, interrupt the infected computer's communication with other machines, erase files, type messages in

a program that the user is currently running, open the CD-ROM drive door, play strange noises through the speaker, and even reboot the computer.

File serving Trojan horse This type of Trojan horse creates a file server (similar to the FTP server) on the infected machine. With this file server, the intruder (attacker) can control network connections, upload, and download files.

The file serving Trojan horse is so small in size (may be not more than 10 Kb) that it is difficult to be detected. It is often hidden in online games, funny forwarded messages, attachments in emails, or in other files that users may download from the Internet.

Denial of service attack Trojan horse This type is usually targeted to a primary server to enable a hacker gain control over one, several, or all computers in its network. After gaining the control, hackers flood the target server with traffic, thereby making it impossible for users to access certain websites.

Keylogging Trojan horse Keyloggers record every step of the user's activity on the infected computer (with regard to the mouse clicks and keys pressed). It emails the recorded information about keystrokes to the hacker. Hackers use this information for performing card fraud and identity theft. For example, hackers get information about username, password, credit card number, pins, and other valuable data to commit online thefts.

Password stealing Trojan horse This type of Trojan horse is used to steal passwords. Like keyloggers, this Trojan also transmits information about passwords to the hacker through email.

System killing Trojan horse These Trojans destroy everything in the system. Examples include Trojan.Killfiles.904 and Trojan.KillAV.

Trojan dropper This type of Trojan horse drops or downloads additional malicious files to the computer. These malicious files further infect the computer.

Joke Trojans Such a Trojan horse causes no damage to the computer but plays an annoying sound from the speaker or displays irrelevant messages on the screen like 'Now formatting hard drive'.

Icandance Trojan This Trojan horse causes no harm to the computer but minimizes all application windows and then starts rapidly scrambling all the desktop icons.

Some more Trojan horses Here are some more types of Trojan horses:

- Rootkit prevents malicious programs from getting detected.
- Trojan-FakeAV simulates the activity of antivirus software and aims at extorting money from innocent users for detecting and removing threats that are not even present on their computer.
- Trojan-Game Thief steals user account information from online gamers.
- Trojan-IM steals user's login details and passwords for instant messaging programs such as ICQ, MSN Messenger, AOL Instant Messenger, Yahoo Pager, and Skype.
- Trojan-Ransom prevents the users from accessing data stored on their computers. The attacker unblocks the data only after users pay them the ransom money demanded.
- Trojan-SMS sends text messages to premium rate phone numbers from the user's mobile phone.
- Trojan-Spy spies the user's activities on the computer using keyloggers by taking screenshots or getting a list of running applications.



Most antivirus software can detect and remove Trojan horses.

4.3.2 Spyware

A spyware is a malicious program that surreptitiously monitors the activity on a computer and reports that information to others without the user's consent. Spyware is usually used for tracking and storing the user's Internet browsing patterns, gaining information about user logins, bank, or credit card information; serving up pop-up ads to Internet users; installing additional software; redirecting web browsers to untrusted sites; modifying software settings; reading cookies; reducing network connection speeds; and causing slowdown or even crashing a computer system.

Spywares are bundled as a hidden piece of code in a freeware or shareware program, which can be difficult to remove once downloaded from the Internet.

4.3.3 Adware

Also known as advertising-supported software, adware is any software that is given to the user with advertisements embedded in the application. When users download a freeware or a shareware from the Internet that has an adware embedded in it, the adware gets installed in the user's computer. They can also spread through email attachments and shared files.

Adware comes under the category of malware because most of the times, they are unwanted. They are a form of spyware that tracks the user's Internet surfing habits and collects information about the user to display advertisements related to them. Users try to avoid them as they see it as a threat to their privacy and security over the Internet and also get annoyed due to the distraction caused by them.



Adware displays advertisements automatically without the user's permission.

Unlike spyware, adware does not transfer the user's personal information to another location. Both, however, slow down the computer speed and allow constant pop-up advertisements to plague the user.

4.4 NETWORK ATTACKS

We will discuss some network attacks in this section.

Denial of service (DoS) attack It is an attempt to make a computer resource unavailable to its intended users. Generally, DoS attacks (Figure 4.2) target high-profile web servers such as banks, credit card payment gateways, government organizations, media, and root name servers. The common ways in which DoS attacks can be made are as follows.

Ping of flood In this scheme, the attacker sends numerous external communication requests to the target machine so that either it is unable to respond to legitimate traffic, or responds so slowly as to be rendered essentially unavailable.

Ping of death Normally, a ping packet in the network should not exceed 65,535 bytes. A packet greater than this is not only difficult to handle but may also cause the system to crash. Therefore, in this type of attack, ping packets of size greater than 65,535 bytes are deliberately sent.

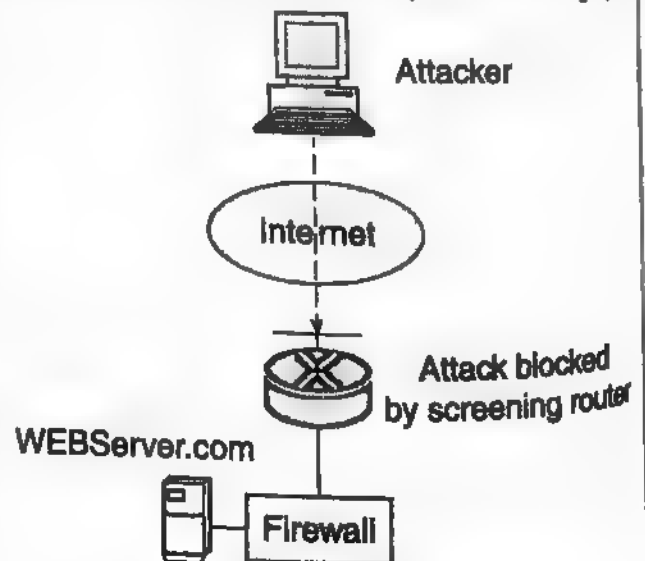


Figure 4.2 Denial of service attack

Teardrop attack In routing, we have learnt that a message is split into packets. Each packet contains a sequence number so that the receiver can reassemble the message accurately. In the teardrop method of DoS, the attacker puts a confusing sequence number in the packets, thereby making it difficult for the receiver to reassemble the message. This may finally result in system crash.

Mail bombs Unauthorized users may send a large number of email messages with large attachments to a particular mail server to fill its disk space so that other users are denied email services.

In other types of DoS attacks, the attacker identifies serious bugs in the target computer system and causes the target system to crash by sending an input that takes advantage of bugs. These bugs may lead to system crash or severely destabilize the system to an extent that it becomes unable to be accessed or used by other users.

4.4.1 Browser Hijacking

When a browser is hijacked, the attacker modifies the browser to permanently change the home page. This is basically done to boost web traffic hits or just to be asinine. Browser hijacking attacks are really irritating and most of the times, even good antivirus software take a long time to detect and free innocent users from such attacks.

4.4.2 Denial-of-Service Attack

This type of attack is usually targeted on a primary server to enable a hacker gain control over one, several, or all computers in its network. After gaining the control, hackers flood the target server with traffic, thereby making it impossible for users to access certain websites.

The most common type of denial-of-service (DoS) attack occurs when an attacker 'floods' a network with information. When we type the URL of a particular web page in the address bar of the browser, the web browser sends a request to that site's computer server to send that page. The web server can only respond to a certain number of requests at once, so if an attacker overloads the server with requests, the web server either crashes or stops responding to the legitimate users.

The attacker can even use spam emails to launch a similar attack on an email account. For example, every email account has a fixed amount of space for storing emails. By sending many or large email messages to the account, an attacker can consume all the storage space to prevent the receipt of any legitimate email message.

Thus, a DoS attack targets the availability of web applications. Unlike other kinds of attacks, DoS attacks are not meant to steal information but they are simply done to slow or take down a website. These attacks can be very problematic, especially when they cause large websites to be unavailable during high-traffic times. Popular websites such as Google, Twitter, and WordPress have been a victim of this attack. However, security softwares have been developed that detect DoS attacks and limit their effectiveness.



Although DoS attacks do not cause loss or theft of confidential data, it can cost the victim a great deal of time and money.

4.4.3 Man-in-the-middle Attack

In a man-in-the-middle (MIM) attack or network spoofing attack, the attacker intentionally inserts himself into a conversation between two persons. Besides deliberately getting in between the persons, the attacker impersonates the people in the conversation (acts as proxy to persons) and gains access to the information that they were sending to each other.

Figure 4.3 shows how an attacker intercepts, sends, and receives data which is meant for someone else. MIM attacks are commonly made for financial gains.

How to Prevent Man-in-the-middle Attacks

The following are the ways to prevent MIM attacks:

- Implement packet filters to inspect packets that are exchanged over the network. Such filters can block packets coming from suspicious IP addresses.

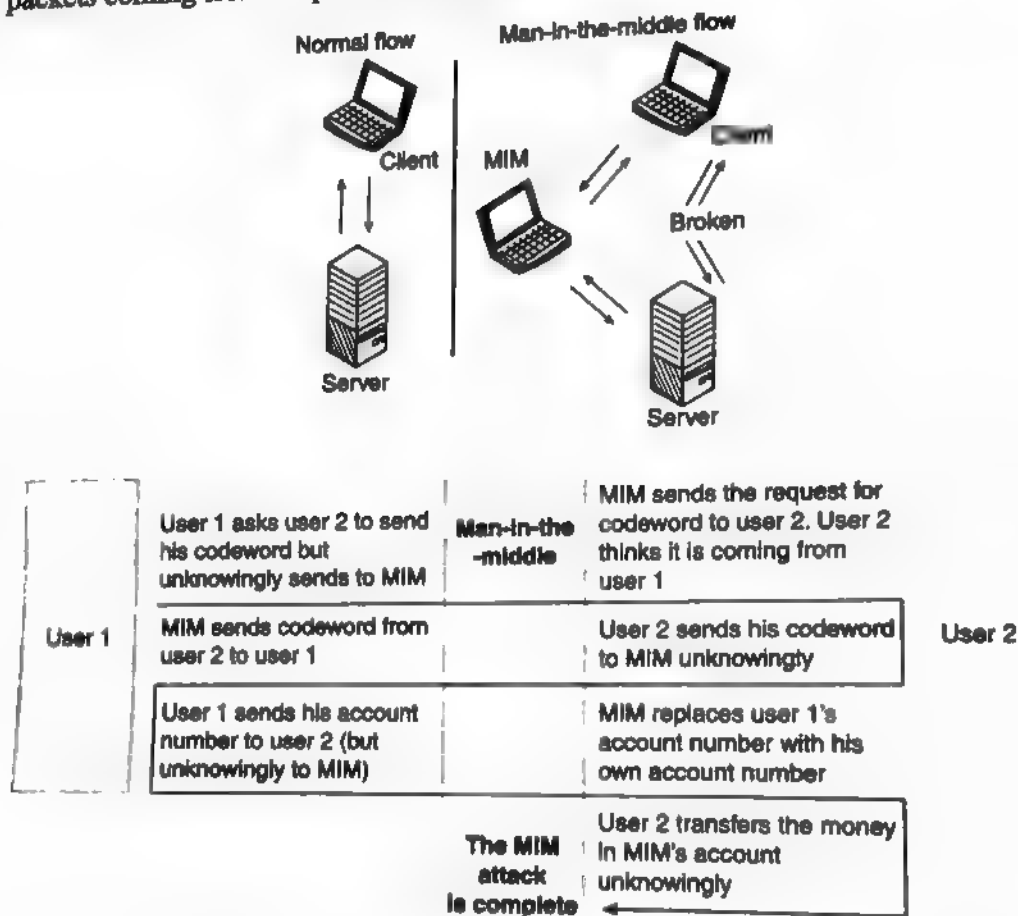


Figure 4.3 MIM attack

- Avoid trust relationships because such relationships only use IP addresses for authentication. The attacker can run spoofing attacks very easily.
- Implement spoofing detection software.
- Use cryptographic network protocols like using HTTPS rather than HTTP. Such protocols send encrypted data and, while receiving, authenticates the data. Other such protocols are transport layer security (TLS) and secure shell (SSH).

4.5 INTERNET FRAUD

The downside of using Internet services include stealing personal information, conducting fraudulent transactions, or transmitting the proceeds of fraud to financial institutions. Such frauds can occur in chat rooms, emails, message boards, or on websites. Some common Internet frauds are discussed here:

- *Purchase fraud* occurs when a criminal purchases a product or service online and pays for it through fraudulent means; for example, using a stolen or a fake credit card. As a result, merchants do not get paid for the transaction and lose money as a result.
- *Online auction fraud* occurs when a fraudster starts an online auction of high-priced items on a website. He accepts payment from the auction winner, but either does not deliver the product or delivers a product that is less valuable than the one offered.
- *Work from home scam* occurs when the scammer accepts services from victims (such as writing directories, data entry, and reading books) but refuses to reimburse them by rejecting their work considering it sub-standard.
- *Phishing* is done to acquire sensitive information such as passwords, account numbers, and credit card details. In this technique, the fraudster constructs a fake website that looks similar to the legitimate site and asks for the user's personal information to steal his information and misuse it.
- *Stock market fraud* includes attempts to manipulate prices of securities on the market for the personal profit of the scammer. For example, the scammer spreads false information to cause a dramatic increase in price of thinly traded stocks and the moment prices reach the desired level, the scammer sells his stocks to innocent victims, thereby making a substantial profit.
- *Online intellectual property theft* is also common these days. Individuals all over the world who share their notes and information on the Internet have exclusive rights on their material. However, many people or students just copy and use it without taking permissions from the author.

4.6 RANSOMWARE

Ransomware is a subset of malware that locks the data (usually by encryption) on the victim's computer and asks for payment to decrypt that data and return access to the victim, as shown in Fig. 4.4. Another variant of ransomware (also known as lock screen ransomware) changes the victim's credentials (username and passwords) so that he/she is unable to use the computing device. The main reason for ransomware attacks is to obtain money from the victim. Unlike other types of attacks, the victim is notified that his/her computer has been exploited and instructions are given on how to recover from the attack. Payment is then demanded in bitcoin (a virtual or digital currency) so that the cybercriminal's identity is not known. The ransom also notifies that if the money is not transferred by a certain date, the key required to unlock the device or decrypt files will be destroyed.



The malicious user sends a bitcoin address to which the bitcoin is to be transferred.

How does Ransomware Spread?

Ransomware malware can be spread by following ways:

- Malicious email attachments
- Spam mails that appear to the victim as if they contain invoice, job offers, security warnings, and other legitimate files
- Infected software apps
- Infected external storage devices
- Compromised websites
- Clicking on pop-up message
- Visiting a website containing malicious advertisements

In many cases, the victim is duped into believing that he/she is the subject of an official inquiry and unlicensed software or illegal web content has been found on his computer. Therefore, he/she needs to pay an electronic fine for the same.

In other cases, the victim is threatened that the data on his computer will be exposed to the general public in its unencrypted state if the money is not paid by a given deadline.



Bitcoin is popular among cybercriminals because it is decentralized, unregulated, and practically impossible to trace.

Is ransomware a virus? No. Viruses infect files or software and have the ability to replicate. Ransomware, on the other hand, scrambles files to make them unusable and then demands money. However, both virus and ransomware can be removed with an antivirus. However, if the files are encrypted, then you may never get them back.

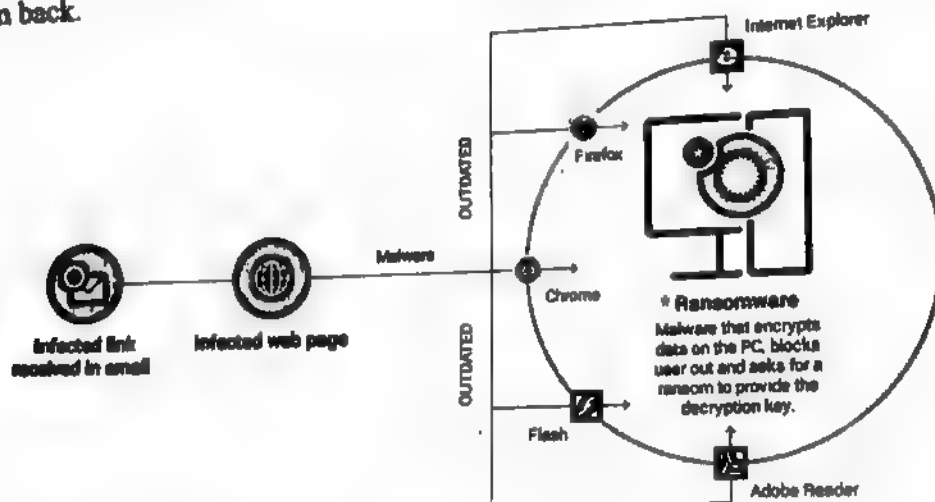


Figure 4.4 Ransomware

Famous Ransomware: CryptoLocker and WannaCry

The first example of a widely spread ransomware attack was Cryptolocker, that was active on the Internet from September 2013 through May 2014. Then, in May 2017, a ransomware called WannaCry infected and encrypted more than a quarter million systems globally. The ransomware had the ability to automatically spread across large networks by exploiting a known bug in Microsoft's Windows operating system.

With this ransomware, the National Health Service in the UK was heavily affected and was forced to effectively take services offline during the attack. According to reports, thousands of companies were affected and incurred losses of more than \$1 billion. It also struck Spanish companies such as Telefonía, Portugal Telecom, the delivery company FedEx, and a Swedish local authority.

WannaCry infected more than 2,00,000 users and 10,000 companies worldwide. Windows XP users were hit hardest by this malware as Microsoft had ended support for this operating system. Microsoft released a patch for the attack when it had become severe.

In 2015, mobile ransomware has also been reported. A malicious Android app called Porn Droid locked the user's phone and changed its access PIN number, demanding a \$500 payment. Although it may seem like a small amount to charge, the ransomware attacks are often widely distributed, so the ransom payments turn out to be a real big amount.

Two major ransomware Locky and Cerber were detected in 2016. Locky was sent to millions of users around the world. It was sent as an email scam which contained an invoice or a receipt of order as a Word file. The victim was asked to enable macros to view its content so that the malware could be downloaded. With every attack, Locky's authors kept on improving the code so that it would be difficult to detect. Cerber malware was also spread via an email attachment or the *Unsubscribe* link in a spam email. It can operate even when the system is offline, and can encrypt more than 400 file types, including database files.

Researchers with the security company Avast had observed 75,000 infections in 99 countries including Russia, Ukraine, and Taiwan.

Security experts have reported a 50% increase in Android ransomware attacks. WannaCry copycat that is spread on gaming forums targets Android devices in China. Since data can easily be restored by syncing devices, cyber criminals often block the smartphone instead of just encrypting files.

Similarly, Mac ransomware is also on the rise. Cyber criminals either target iCloud accounts or try to lock smartphones through the Find My iPhone system.

Ransomware Attacks in 2017

Figure 4.5 shows the 10 countries most affected by WannaCry in 2017.

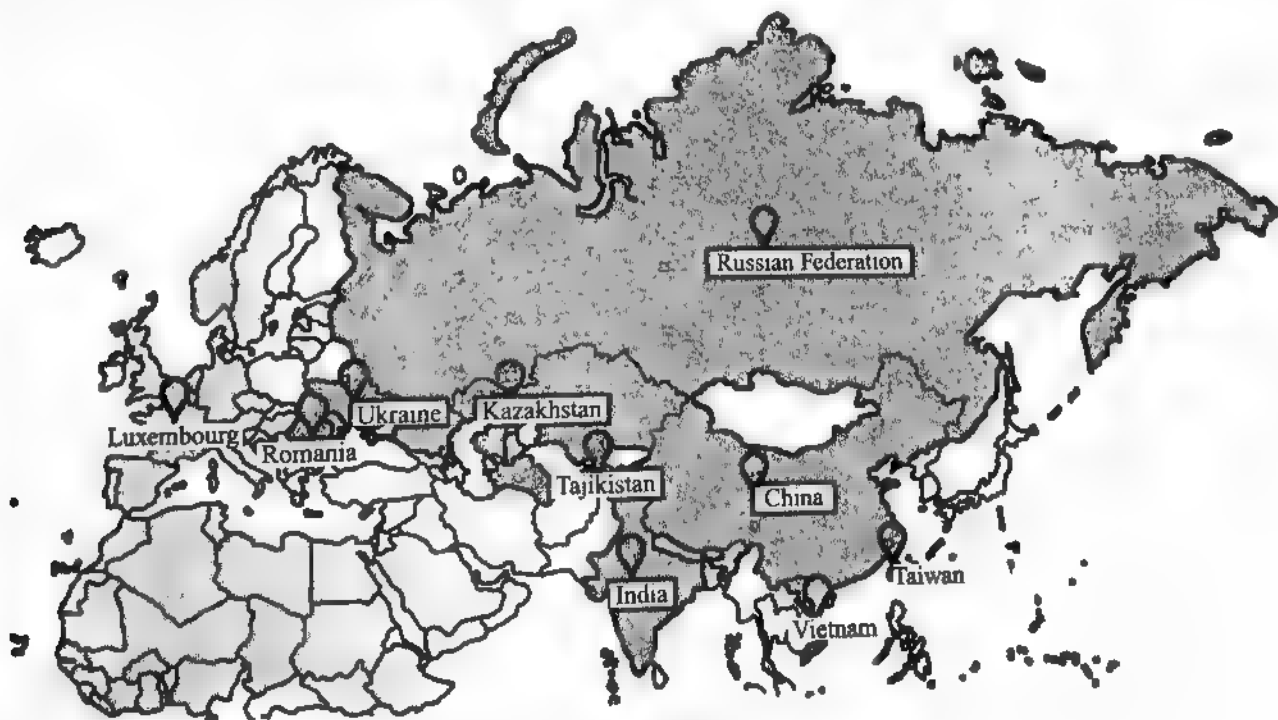


Figure 4.5 Countries affected by ransomware

Ransomware Prevention

While ransomware attacks may be nearly impossible to stop, users can take some measures to ensure that the damage is minimal and recovery is as quick as possible. These measures include the following:

- Taking back up of computing devices on a regular basis
- Using good antivirus programs that scan files before they are downloaded
- Blocking secret installations and looking for malware that may already be on a computer
- Updating software regularly including antivirus software and the operating system

- Being extra cautious while clicking on links in emails from strangers or opening email attachments
- Using proper authorization and access control mechanism
- Installing a pop-up blocker on your computing device
- Using some decryption tools. They may succeed to decrypt and recover some information
- Never paying ransoms because doing so will further encourage them. Moreover computer security experts have warned that there is no guarantee that the access would be granted after even making the payment



If you are using Internet Explorer then turn on the smart screen. It will help you identify reported phishing and malware websites.

Removing ransomware is quite simple as long as it has not locked your computing device. In this case, you can enter Safe Mode and then either remove it manually or use an antivirus to delete it.

However, if the computer is infected with a locker ransomware then you cannot even enter Windows and run applications. To fix this, do a System Restore to restore Windows back to a point in time where your computer was still safe. Second, run the antivirus program from a bootable disk or an external drive. Third, reinstall your operating system.



Locker ransomware locks the victim out of the operating system making it difficult to access the desktop and any application or file. Some of these ransomware infect the Master Boot Record so that the computer is unable to boot up.

4.7 HACKING

Hacking is a practice of identifying weakness or loopholes in a computer system, or a computer application, or a network to **exploit its weaknesses to gain access (of course, unauthorized)**. For example, cracking a password to access a computer is an act of hacking. Therefore, we can say that hacking means gaining unauthorized access for some illicit purpose.

Correspondingly, the person who performs hacking is known as a hacker. There are three types of hackers or we can say that hacking can be done for three reasons.



Do you know that breaking a security system requires more intelligence and expertise than actually creating one? So a hacker is obviously a very intelligent and a highly skilled person in computers.

White hats: A white hat hacker hacks to check the security systems to make it more hack-proof. Mostly, white hackers are a part of the organization whose security they are checking.

Black hat hackers: They hack to take control over the system for personal gains. After gaining access, they may destroy, steal, or even keep authorized users from using the system.

Grey hat hackers: They include all the curious people who use their computer skills to hack a system to locate potential loopholes in the network security system. They then inform the network administrators about the weaknesses so that they can be rectified.

4.8 SPOOFING

The word *spoof* means to hoax, trick, or deceive. Correspondingly, spoofing means tricking or deceiving computer systems or other computer users. This is either done by hiding one's identity or faking the identity of another user on the Internet.

Spoofing can be done in several ways. In the first variant that is **email spoofing**, messages are either sent using a bogus e-mail address or faking the e-mail address of another user. However, since most of the email servers today have security features, it is extremely difficult for unauthorized users to send messages. To counterpart this tight security, spammers often send spam messages using fake e-mail addresses (using their own SMTP).

In the second variant of spoofing, which is also known as **IP spoofing**, the IP address of a certain computer is masked so that it becomes difficult for other systems to determine where the computer is transmitting data from. Since IP spoofing makes it difficult to track the source of a message, it is often used in denial-of-service attacks that overloads a server. This may cause the server to crash or make it unresponsive to legitimate requests. Fortunately, software security systems have been developed that can identify denial-of-service attacks and block their transmissions.

In its third variant, the cyber criminal **fakes an identity** (like an online username). For example, when posting on a web discussion forum, a user may pretend to be a representative of a company, when actually he/she has no association with it. Moreover, in online chat rooms, users may fake their age, gender, and location.

The only way to prevent spoofing is to make sure that you know who you are communicating with. This is especially more important when you are giving out private information over the Internet. Though spoofing may be a part of man-in-the-middle attack, but it is more general.

IP spoofing attacks usually take place when trust relationships use IP addresses (rather than user logins) to verify machines' identities to access systems. Cyber criminals use spoofing attacks to impersonate machines with access permissions.

Prevention

- Avoid trust relationships
- Use spoofing detection software
- Use cryptographic network protocols such as Secure Shell (SSH), HTTP which encrypt data before it is sent and authenticate data as it is received

4.9 SNIFFING

Sniffing means monitoring and capturing data that is being transmitted over a network. While a network professional can do sniffing to diagnose network issues, troubleshooting, and analyse network usage; malicious users, on the other hand, use it to steal usernames and passwords, network details, credit card numbers, etc., to cause monetary and reputational damages. Just imagine how easy it would become for a cyber criminal to access a computer or network when he has all legitimate usernames and passwords with him.

Sniffing can be done using various tools. Sniffing is often known as wiretapping applied to computer networks because it can be compared with 'tapping phone wires' to know about the conversation. Sniffing can be as simple as plugging into the network using a cable or connecting wirelessly to the network and sniffing the entire traffic (whether protected or unprotected). Sniffing is usually done to

gather information that can be used for further attacks or to cause other issues for the network or system owner.

Sniffing tools are not dumb utilities that allow you to view only live traffic, but they also allow sniffers to save the data, analyse, and review it whenever time allows.

Besides software based sniffing tools, there are also hardware protocol analyser devices that plug into the network and can monitor traffic without manipulating it.

Sniffing can be done at any level. It can be done either by an employee of the company or an outsider. It is pretty easy for an internal employee as he is already connected to the network. An external attacker may intercept data at the firewall level or use wireless networks which has made it easy to sit near the network and penetrate it to get information (Fig. 4.6).

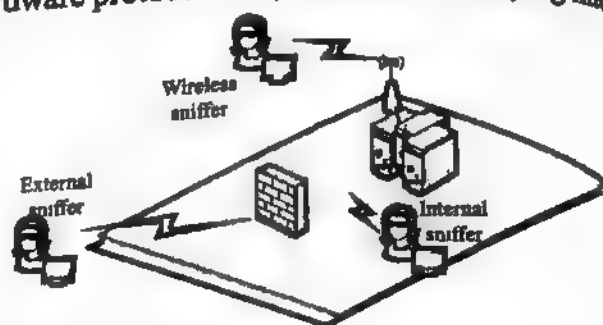


Figure 4.6 Sniffing

Types of Sniffing

Sniffing can be either active or passive in nature.

Passive sniffing In passive sniffing, the sniffer just monitors and captures the data. They do not alter it in any way. So you can say that passive sniffing allows listening only.

Active sniffing In active sniffing, the traffic is not only monitored and captured, but it may also be altered in some way.

Usually, passive sniffing is more difficult to detect and hence a more dangerous attack as the attacker is silent/invisible on the network.

Prevention

To protect confidential information from sniffing, one must take the following measures.

- Use encrypted protocols (such as SSL, HTTPs, SFTP)
- Encrypt all sensitive data before transmitting it. For example, emails can be encrypted using PGP
- Anti-sniffer software (like Wireshark, Sniffit, DSNIFF, Ettercap) can be used to detect known sniffers, processes, etc
- Modern antivirus or anti-spyware software can also be used to detect and then disable sniffing software
- Disabling promiscuous mode on network interfaces which results in shutting down most sniffer softwares

4.10 PHISHING

Phishing is done to acquire sensitive information such as passwords, account numbers, and credit card details. In this technique, the fraudster constructs a fake website that looks similar to the legitimate website. These websites are specifically designed to collect an individual's online bank, credit card, or other login information. Since they seem legitimate web pages, users enter their personal information

thinking they are providing it to their trusted company. Phishing can also be done by sending emails to target the innocent users.

Identifying a Phishing E-mail

Certain emails are sent out to thousands of different e-mail addresses. Most of the times, the phishers do not even know to whom the email addresses belong. Therefore, it is advised that one should not open an email from a company that one is not dealing with. For example, do not believe in emails coming from a bank in which you do not even have an account.

There may be a small spelling and grammar mistake in the fake email. For example, an email supposedly coming from Facebook may spell Facebook as Faceebook. Hence, one should check for obvious errors.

If the email says that there is an error in your account details and you need to rectify them by giving the details afresh by replying to the email, then extra care should be taken to verify the authenticity of the mail.

E-mails that request an immediate response or give a specific deadline may be fake. For example, if the email asks you to click on the link to log in and change your account information within 24 hours, it may not be a genuine one.

Phishing emails often have fake links. They often ask you to click on links that seem to, but do not actually belong to the original company. For example, "<http://fakeaddress.com/facebook>" is not a Facebook URL. It just specifies a URL with a "Facebook" directory. Another example could be replacing 'r' with 'n' or 'm' with 'n' in a big URL so that the fraud goes unidentified as in <http://www.bank-ofindia.com/>.

4.11 SPAMMING

Spam emails are a common form of fraud in which the fraudster sends bulk emails to millions of email addresses to corrupt the receivers' computers, steal their identity, or fool them to pay for fraudulent products or services. These emails offer false dealings to recipients such as low-interest loans, winning lotteries, fancy business proposals, free credit report checks, and relationships with local singles. Spam emails require recipients to open the email and click on a link which may also open up the computer to a virus, worm, or other bug that will corrupt the computer.

An email having an attachment from an unknown sender or from a known sender from whom no attachment was expected should not be opened as it may contain malicious code. Only .txt file type is always safe to click.

Therefore, we can say that *spamming* is the process of flooding the Internet with many copies of the same message for commercial advertising, usually for dubious products such as get-rich schemes or loans at low interest rates.

Spam can be not only annoying, but also dangerous to consumers. Such emails are simply too good to be true. They give lucrative offers and eye-catching or attention-grabbing statements to attract people's attention immediately. For instance, certain emails states that you have won an iPhoneX, a lottery, or a prize, or have been given free/easy loan, etc.

Though, these emails are not requested by a user, they are sent. The source and identity of the sender is mostly anonymous or unknown.

4.12 TECHNOLOGY SOLUTIONS

The entire concept of data security rests on the CIA (Confidentiality, Integrity, and Availability) principle. In this context, confidentiality refers to the ability to hide information from unauthorized people. To ensure confidentiality, data is encrypted before it is transferred from one computer to another. In this section, we will read about different ways of encrypting data.



Integrity ensures that data is accurate and **availability** means that the information is always accessible to authorized users.

4.12.1 Data Encryption and Decryption

Encryption (Figure 4.7) is the process of converting data into a cipher text (random data which is meaningless). Here, cipher text means scrambled data, which cannot be easily understood by anyone except authorized parties. For example, if I want to send a message HELLO to my friend and I do not want anyone else to read my message, then I can encrypt this message and send the encrypted text across the network. The corresponding cipher text for HELLO could be KHOOR. KHOOR is not understood by anyone.

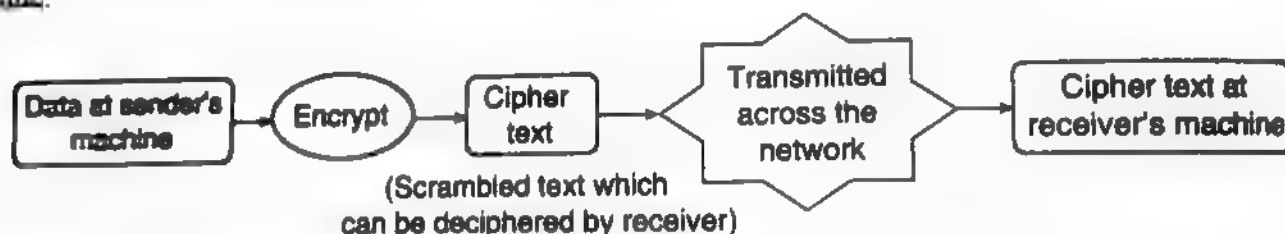


Figure 4.7 Data encryption

Advantage Encryption is done to protect the confidentiality of digital data that is either stored on the computer or is transferred across the network. Many companies store encrypted data in their databases to ensure that even if an attacker gets illegal access to the confidential data, then at least he is not able to understand it.

Nowadays, many encryption algorithms are used; besides providing confidentiality, a sound encryption algorithm has the following features:

- **Authentication** to verify the originator of the message.
- **Integrity** to ensure that the message has not been modified or tampered during transmission.
- **Non-repudiation** to make sure that the sender cannot deny sending the message.

Decryption is the reverse of encryption. It is the process of converting encrypted data back into its original form so that receiver can correctly interpret its meaning. Until decrypted, the cipher text appears as garbage. To protect data from being decrypted by just anybody on the network, only those users who have the decryption key (3 in our example) can decrypt the data and make it useful. For example, only the intended receiver knows that while encrypting the data, the third character (in the forward direction) was replaced by the original character. Therefore, to decrypt the message the third character (in the backward direction) is written.



If the decryption key is not known, even then the attacker may decrypt the message by applying several decryption algorithms.

4.12.2 Symmetric and Asymmetric Encryption

Encryption is done using two main techniques (Figure 4.8)—symmetric (also known as secret-key, single-key, shared-key, one-key, or private-key) encryption and asymmetric (public key) encryption.

Symmetric Encryption

This type of encryption is basically done on small amount of data. It uses a symmetric key which is applied on plain text (data) to convert it into cipher text. Similarly, during decryption, the symmetric key is applied to convert the cipher text into original data.

Strength A good symmetric encryption algorithm is one that makes it very difficult, if not impossible, for attackers to decrypt the generated cipher text without knowing the key used for encryption.

Requirements The following are the requirements of systematic encryption:

- The longer the key, the more difficult it will be to decrypt the message. Most symmetric encryption algorithms use a key of 64 bits, 128 bits, 256 bits, and even 512 bits.
- It is always better to use an encryption algorithm that has been used for several years and has successfully resisted all attacks.
- The secret key can be a number, a word, or just a string of random letters. For example, 'hello' can be encrypted as 'hzexlqlhog' by inserting a character randomly between any two characters in the original message.
- Both the sender and the receiver know about the secret key.
- The secret key must be changed on a regular basis.

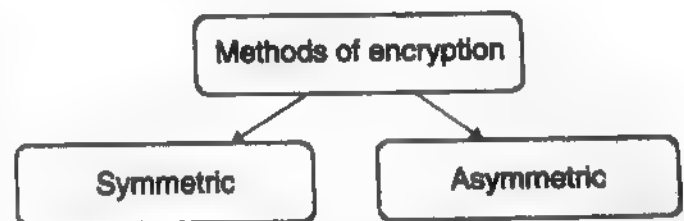


Figure 4.8 Encryption methods

Drawbacks The following are the drawbacks of systematic encryption:

- While exchanging the secret key with the receiver across the network, it may get into the hands of an attacker. Once the key is known to him, he can decrypt all the messages very easily, thereby defeating the whole idea of ensuring data confidentiality.
- There is no provision for authenticating the sender. There is no way for the receiver to know whether the message has been sent by the intended sender.
- Data integrity cannot be assured—whether the received message is the one that was sent.

Asymmetric Encryption

Asymmetric encryption algorithm overcomes the limitations of symmetric encryption algorithm by using a pair of keys. The two keys in asymmetric encryption are related to each other in such a way that a message encrypted by one key can be decrypted only by the second key.

The two keys are known as public key and private key. While the public key is given to anyone who wants to send a message, the private key on the other hand is kept secret and is known only to its owner. **Technique** A message (which may include a small text or a file) can be encrypted by using the public key. Similarly, any message encrypted by using the public key can be decrypted only by applying the equivalent private key on the cipher text. This process is shown in Figure 4.9.

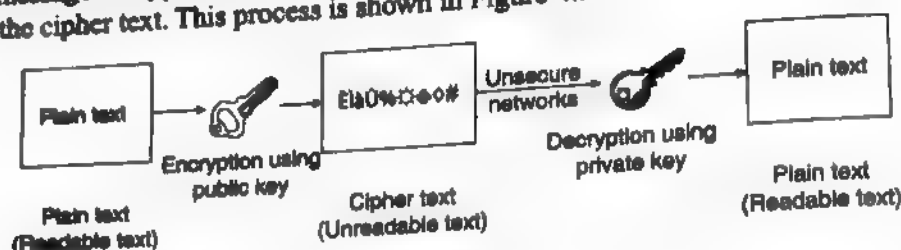


Figure 4.9 Asymmetric encryption



Public and private keys are allotted by a certificate authority.

As per Figure 4.9, the sender uses the receiver's public key to encrypt the message. The encrypted text then travels across the network. The receiver decrypts the cipher text with his private key to read the confidential message.

Example If my public key is P1 and my private key is P2, then the key P1 is known to everyone who wants to exchange data with me. Only I know my private key P2. A user X who wants to send me a confidential data will encrypt the data using P1. The data encrypted by P1 can be transformed into original text only by applying P2. This ensures that the message intended to be read by me is actually read only by me.

Key features The key features of asymmetric encryption are as follows:

- The two keys are mathematically related to each other.
- The algorithm used for encryption is universally known.
- It is impossible to compute the private key if the public key is known.

Table 4.1 differentiates between symmetric and asymmetric encryption techniques.

Table 4.1 Differences between symmetric and asymmetric encryption

Symmetric encryption	Asymmetric encryption
A single secret key is used.	A pair of keys is used.
Secret key is known to sender and receiver.	Public key is known but private key is kept secret.
It is used by the Digital Encryption Standard.	It is used by Pretty Good Privacy.
It is computationally faster.	It is slower than symmetric encryption.
It is less complicated.	It is more complex.
Secret key is shared.	Public key is available to everyone; private key is kept secret; and no key is shared.

4.13 SECURITY

Today, we are living in a world that is parallel to the walls, DMZs, and firewalls.

4.13.1 FIREWALLS

You must have heard that a website is the presence of a company in the internet and goes in and goes out of the company's network.

The firewall is a security system that is designed to prevent unauthorized access to or from a private network.

Firewalls are used to protect every computer, Linux, testing, turned, users of computers, Zone, Ice Pr, Micro, free of charge.



Types of Firewalls

The types of firewalls are:

- Packet Filtering Firewall
- Stateful Inspection Firewall
- Proxy Firewall
- Dynamic Packet Filtering Firewall
- Static Packet Filtering Firewall
- Stateful Packet Filtering Firewall
- Next-Generation Firewall

4.13 SECURITY IMPLEMENTATION

Today, we cannot rely on just one way of securing our data. Other techniques should be implemented in parallel to protect our computer or network systems. In this section, we will therefore read about firewalls, DMZ, HTTPs, SSL, and website auditing.

4.13.1 Firewalls

You must have heard people saying, 'We cannot access Facebook in our office.' How is it possible that a website available all over world cannot be accessed in a particular location? The answer to this question is the presence of firewall. A firewall that may be a piece of hardware, software, or both is installed to prevent unauthorized access to computers or networks. It exercises full control over data packets coming in and going out of the computer to the Internet. All data traffic passes through the firewall as shown in Figure 4.10.

The firewalls are configured with a set of rules that decide the packets to be accepted that will be transmitted or received across the network. Although firewalls are already programmed with a sensible set of rules by the manufacturing company, users can also configure them. Basically, a firewall is designed to act as a barrier to keep destructive or unwanted forces away from your property.

Firewalls have become so important today that every operating system including Windows, Mac, Linux, etc. offer built-in support for maintaining and testing firewalls on the computer and the firewall is turned on by default. Besides the built-in support, users can also use third-party firewalls to protect their computers. Some examples of third-party tools are Zone Alarm, Norton Personal Firewall, Tiny, Black Ice Protection, McAfee Personal Firewall, and Trend Micro PC-cillin. Many of these tools either offer a free or a trial version of their commercial versions.

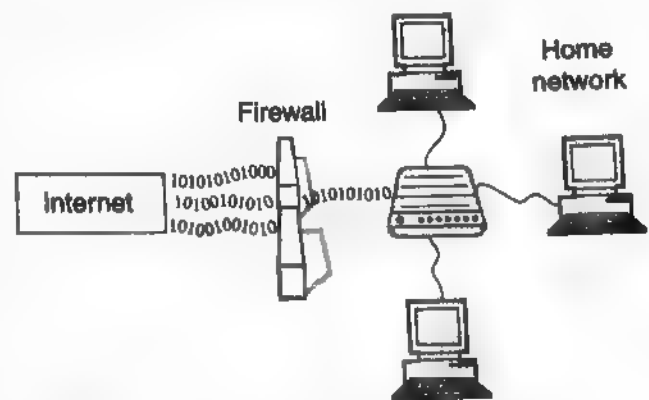


Figure 4.10 Firewall



Firewall helps to screen out hackers and malicious software that tries to reach a computer.

Types of Firewalls

The different types of firewalls are explained in this section.

Packet filtering This technique examines each data packet that either enters or leaves the network. Based on the result of examination, it accepts or rejects the packet depending on user-defined rules. Packet filters work mainly on the first three layers of the OSI model. Although packet filtering is a fast, efficient, and effective technique, it is difficult to configure and is susceptible to spoofing. Moreover, packet filter firewalls cannot tell whether a packet is part of an existing data exchange or a new one. This is because each packet is treated in isolation.

Stateful firewalls Stateful firewall overcomes the drawback of packet filters by recording all connections passing through it. This gives enough information to determine whether a packet is the start of a new connection, an existing connection, or not a part of any connection. When the state of the packet

is known, the firewall can speed up packet processing by allowing a packet of an existing connection without further analysis and evaluating only those packets (based on rules) that are coming through a new connection.

Circuit-level gateway implementation In this type of firewall, security mechanisms are implemented only when a TCP or UDP connection is established. Once the connection is established, packets can flow between the hosts without further checking.



A proxy service must be implemented for each type of Internet application the firewall will support. For example, there is an HTTP proxy for web services.

Proxy firewalls Proxy firewalls act as an intermediary for requests from one network to another to disallow any direct connections between either sides of the firewall. This enables the proxy firewall to block or allow traffic based on its rule set. A proxy service must be run for each type of Internet application the firewall will support, like an HTTP proxy for web services.

Application-layer firewalls With increasing attacks against web servers, a firewall is required to protect servers and the applications running on them. For this, application firewalls are implemented that inspect and filter packets on any OSI layer (up to the application layer). Such a firewall can block specific content (such as malware or certain websites) and report when applications and protocols (such as HTTP, FTP, and DNS) are misused.

However, practically speaking, most firewalls use more than one technique to implement security mechanisms.

Firewall Rules

Firewalls can be configured based on user-defined rules. These rules can be based on the following features.

IP addresses Block data coming from or going to a certain IP address or a range of IP addresses.

Domain names Allow or disallow data from certain specific domain names or domain name extensions such as .edu or .mil.

Protocols Allow or disallow data that uses protocols such as IP, SMTP, FTP, UDP, ICMP, and Telnet.

Keywords Allow or disallow data flow that contains certain keywords or phrases. This is done to block offensive or unwanted data from flowing in.

4.13.2 Demilitarized Zone (DMZ)

The term DMZ was borrowed from military terminology. Basically, a DMZ is an area between two territories that are hostile to one another. This term was first used to refer to the strip of land that separates North Korea from South Korea. Similarly, in computer networks, DMZ (also known as perimeter network) separates an internal network from the often hostile territory of the Internet.

In organizations, computers that need to communicate directly with the Internet (or any other public servers) are placed in the DMZ and not in the internal network. Although these computers are protected by the firewall, they are still at risk as they have direct contact with Internet computers. The DMZ is only 'semi-secure'. It is easier to hack a computer in the DMZ than on the internal network. So even if a computer in DMZ is hacked, the security of the internal network is not compromised because it is on a completely separate, isolated network.

DMZ provides an additional layer of security to the LAN as it restricts hackers from directly accessing internal servers and data via the Internet. A DMZ can also be set up on home networks, although its usefulness in homes is limited.

Note that DMZ can be a logical or a physical network acting as a secure bridge between an internal and external network. It has limited access to the internal network, and all of its communication is scanned on a firewall before being transferred internally.



DMZ is considered more secure, safer than a firewall, and can also work as a proxy server.

4.13.3 Secure Socket Layer

Secure socket layer (SSL) is a security protocol that uses encryption of messages exchanged between a web server and a browser in an online communication. To create an SSL connection, one must have a SSL certificate. To get this certificate, one must follow the following steps.

Step 1: Give all details about his/her website and company.

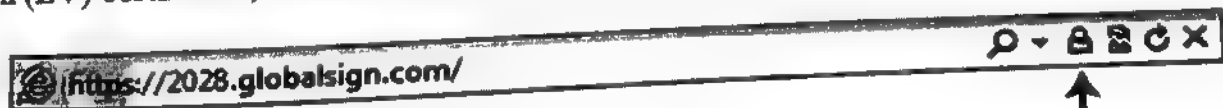
Step 2: A pair of private and public key is then assigned to the applicant.

Step 3: Certificate signing request (CSR) is then submitted to the certification authority (CA). CSR is a data file that contains applicant's details as well as the public key assigned to him/ her.

Step 4: The CA validates the details.

Step 5: If the authentication is successful, SSL certificate is issued. This certificate would be matched to the private key. The certificate includes the domain name, the name of the company, expiration date of SSL, details of CA and other things like applicant's address, city, state, and country. Now, an encrypted link is established. The existence of an encrypted session is indicated by the presence of the lock icon in the address bar.

Note that when SSL certificate is installed on a web server, a secure connection is established between the web server and the browser that connects to it (Fig. 4.11). Moreover, the website's URL is prefixed with "https" instead of "http" and a padlock is shown on the address bar. If the website uses an extended validation (EV) certificate, then the browser has a green address bar.



The standard HTTP is changed to HTTPS, automatically telling the browser that the connection between the server and browser must be secured using SSL.

The padlock is activated, showing you that the browser connection to the server is now secure. If there is no padlock or the padlock shows a broken symbol, the page does not use SSL.

Figure 4.11 SSL protocol

Use of SSL

The SSL protocol is used by millions of online businesses to protect their customer's confidential data which may include their personal information, passwords, or credit card details (Fig. 4.12). All web browsers can interact with secured websites as long as the site's certificate is issued by a trusted CA.

SSL is even more important in today's scenario when fraudsters and cyber criminals do not even leave a single opportunity to steal valuable data including bank account numbers and card details. Any moderately skilled hacker can easily intercept and read the data being transmitted. Therefore, the need for an encrypted link cannot be ignored.

Moreover, SSL is the most widely used security protocol today and, according to Google, it is being used to secure more than 50% of the pages loaded by the Chrome browser. Besides using this protocol for transmission of web pages, SSL is also used for applications including email, file transfer, instant messaging, and voice over IP.

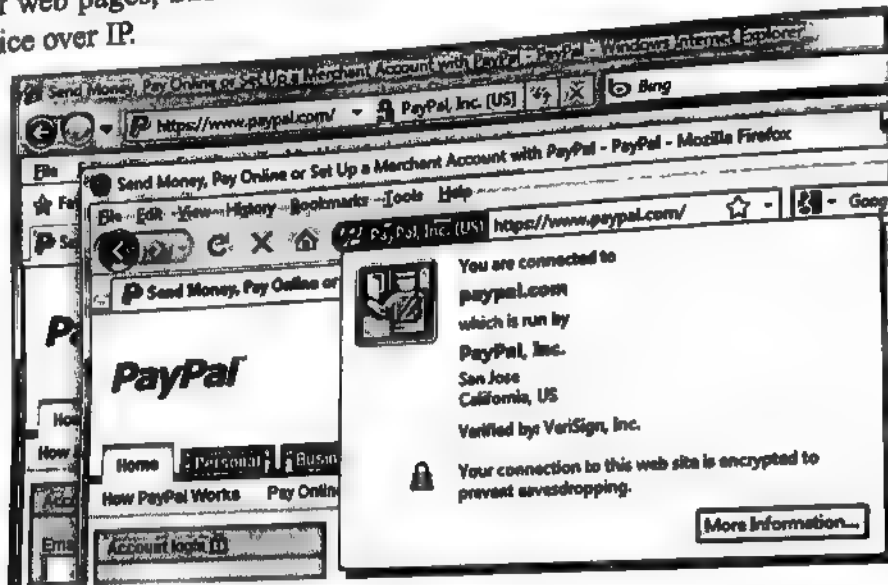


Figure 4.12 SSL protocol in action

Establishing an SSL-encrypted Connection

Whenever a browser initiates a connection with a SSL secured website, the following things happen:

- A user requests a secure connection to a website by entering a URL that begins with 'HTTPS' in the address bar.
- The browser will first retrieve the site's SSL certificate to check if it is still valid (that it has not expired or been revoked).
- It then checks whether the CA is trusted by the browser.
- Then another check is made to ensure that the certificate is being used by the website for which it has been issued.
- If all the above checks are successful, then the communicating parties exchange data over the encrypted link. In case of failure, a warning is displayed to the user, indicating that the website is not secured by a valid SSL certificate.
- Once an encrypted connection is established, data transferred between the two parties can only be deciphered by the two parties involved. Even if hackers get that data, they will not be able to decode it.

How Does SSL Work?

We discussed that whenever a web browser initiates data exchange over a secured link, it first makes a lot of checks. In addition to the aforementioned checks, additional checks are also made to ensure whether the website confirms to the required security standards on key lengths and other items or if the domain listed on the certificate matches the domain that was requested by the user.

- When the browser confirms that the website can be trusted, it creates a symmetric session key that it encrypts with the public key in the website's certificate. The session key is then sent to the web server.
- The web server uses its private key to decrypt the symmetric session key.
- The server sends back an acknowledgement that is encrypted with the session key.
- From now on, all data transmitted between the server and the browser is encrypted and secure.

Advantages of SSL

- Data is protected using encryption.
- Website can build trust with users and gain more number of clicks and business.
- This means your website visitors will be able to clearly see you have a secure operation and will be more inclined to complete purchases with you.
- Users are sure about the identity of the website. So they cannot be fooled by fraudsters.
- Various search engines give (search engine optimization) benefits to https pages over non-https pages. Increase in SEO ranking increases customer base of the website.
- It protects against phishing attacks.



It is very difficult for cyber criminals to receive a proper SSL certificate.

Disadvantages of SSL

A lot of cost is involved to set up a trusted infrastructure and validate the identity. Performance degrades with SSL as more server resources are required for encryption and decryption of data.

4.13.4 HTTPs

HTTPs is a secure version of HyperText Transfer Protocol (HTTP). The 's' stands for secure. HTTPs ensures secure communication between a user's browser and a web server. A website supporting HTTPs protocol has its URL beginning with https://. Moreover, whenever you are accessing a secured website, you can always see a green address bar or padlock in the browser window.

HTTPs is very much important for e-commerce websites and those websites that accept online payments or confidential data. It sends this data securely to the server to prevent it from being stolen by malicious users. Besides these reasons, HTTPs has also become important because now Google has announced that HTTPS will be a factor in their ranking of the websites. HTTPs either uses SSL or TLS (Transport Layer Security) for encrypting data.

To summarize, benefits of using HTTPs over HTTP include the following:

- Enhances trust of customers by reassuring them that they are using a secure and responsible business
- Customers are sure about the identity of the business
- Increases the ranking by search engine

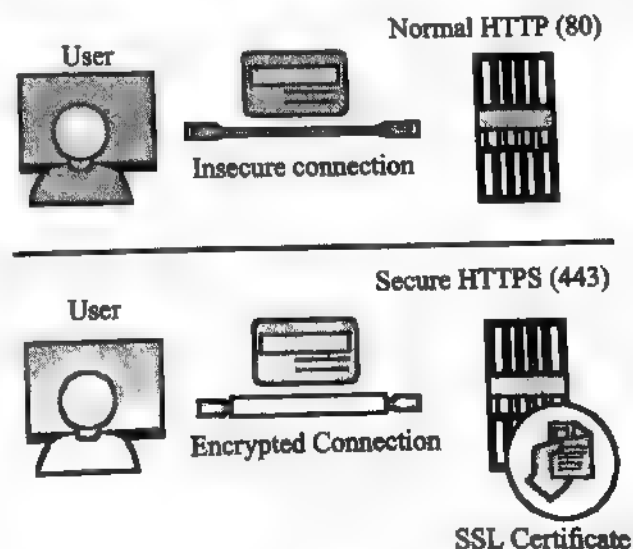


Figure 4.13 HTTP vs HTTPs

- Besides providing confidentiality, HTTPS also provides authentication. To use HTTPS, a website needs to authenticate itself to prove that it's the site it says it is. For this, the website's administrator asks a CA (certification authority such as Comodo or Symantec) to issue the site a certificate that includes a cryptographic key that in theory cannot be forged.
- HTTPS prevents tampering of data by ISPs or hackers. They can no longer insert ads or inject code designed to compromise a user's computer.

Without HTTPS, a government censor can choose to block certain pages of a site or even just parts of a page. Figure 4.13 shows a comparison of HTTP and HTTPS.

HTTP vs HTTPS

On a website supporting HTTP, confidential data can be intercepted, spied on and even altered by anyone between the user's website and the web server.

When a HTTP website is accessed, the web server responds to user's requests (from the browser). All the data is exchanged in an unencrypted form. On the other hand, with a HTTPS website, the user's web browser and the server first exchange cryptographic keys. These keys are then used by the web browser and the web server to send messages that only the other side can decrypt, thereby locking out all hackers.

HTTPS also provides additional privacy for normal web browsing, making your normal web browsing also confidential. So any other person will not be able to know the pages you visited.

Figure 4.14 (a) shows a HTTPS website.

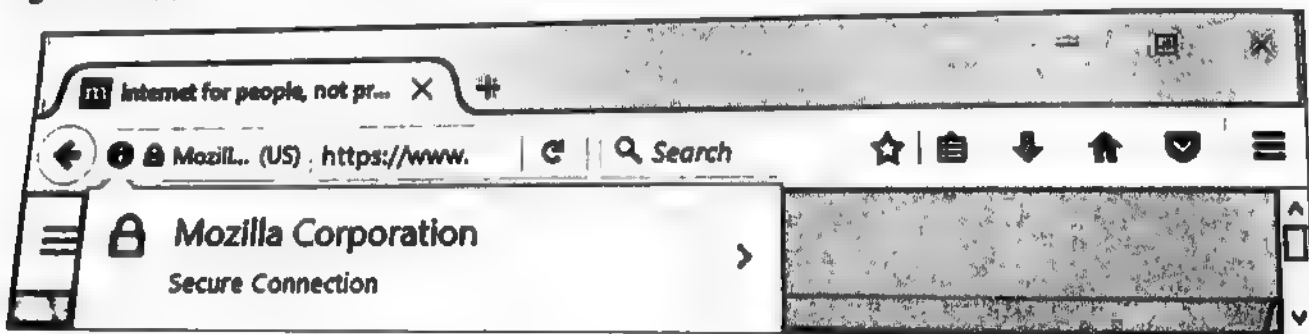


Figure 4.14 (a) HTTPS website

Note if you click on the lock you can see the details of the website as shown in Fig. 4.14(b).



Figure 4.14 (b) Security details of the HTTPS website

4.14 WEBSITE AUDIT

A website audit includes all activities to review the forms, information, content, graphics, evaluation of ranking, and the quality of a website. It is especially done to identify and improve any shortcomings that might have been otherwise ignored. During a website audit, factors like traffic are also monitored as it plays a key role in the success of a website and generation of sales.

Significance of Website Audit

- It gives an insight into new graphic and strategic ideas. For example, with statistics such as bounce rates, time spent on the website, and number of views, a company can discover what content their audience enjoys and finds helpful and what turns them off. This information also helps business to improve both blog and social media strategies.
- It helps to identify elements that are hampering the website's current performance.
- It reveals a website's traffic ranking, bounce rates, and search percentages. The traffic report indicates where traffic to the site is coming from, which search terms (keywords) are being used to trigger the site and which websites provide referrals.
- It helps to create new online marketing strategies to bring traffic to the website.
- A complete website audit reveals broken links and other errors in the HTML code. For example, all 404 errors need to be either rectified or redirected.
- It helps to review response time and download time. This is very important as a slow loading website can hinder the success of the website.
- Character setting is checked as wrong outputs of a page can affect the indexing of website by search engines.
- The file size of the website is checked as a large file size increases the download time.
- The page titles and content of the website are checked to ensure that there is no duplication within the site.
- Know about the most important keywords that are driving traffic to the website.
- It helps to know about areas for optimization in content. For example, content that is irrelevant or outdated and needs to be removed or the content that needs to be redirected.
- It also helps to know which images are not properly tagged.
- Checks metadata to ensure that all pages have unique and relevant metadata. Metadata and descriptions of webpages if not given properly will make it difficult for search engines to find them as they would not be able to make out clearly what the web page or website is all about.
- Check that heading size is uniform and ok (neither too small nor too big).
- Ensure that the website functions properly on mobiles and tablets.
- Ensure all web pages have proper URL structures.
- Checks whether all important pages on the website are properly indexed and those that need not be indexed are blocked (for example, a 'thank you' page).

Web audit saves the website from penalization. If the search engine finds that the web page title or the keywords are not relevant as per its contents, then the website will be penalized. For example, if there have been instances of keyword stuffing or link farming in the website to improve its ranking and traffic, then the website will have to bear penalty from the search engines. A sound website audit can identify and correct anything that can lead to instances penalization.

It also helps website to discover malware infection. When a website is infected, the search engines will come to know about it instantly. So the search engine would display the website in search results but along with a warning to users that they may get a virus if they visit the site. This will automatically bring down the count of visitors to the website to zero and will hamper the reputation of the company. A website audit, therefore is necessary to detect and delete malware before problems become severe.



Search engine optimization (SEO) is used to optimize the website; a website audit, on the other hand, will monitor its optimizing process.

Hence, a website audit helps the company to increase traffic and page rankings and thus increase the return on investment (ROI). In short, it provides everything one needs to know about monitoring and maintaining the website. Website audit should be an ongoing process that should be conducted at least once a year.



There are many websites on the Internet that have automated SEO audit tools which can be used by anyone. You just need to enter the URL of the website and the report will be generated in a few minutes.

Points to consider before going for a Web Audit

- The website should be organized in such a way that users can easily navigate through it.
- There should not be too many items in the navigation bar. Instead of cluttering the navigation bar, dropdowns should be used to list all the services.
- There should be minimum number of tabs in the main menu.
- Make sure that users do not have to click more than three times to open a particular page.
- Every page should have a link, button or any other way to return to the homepage.
- The homepage should not have excess of information.
- There should be a utility to search for the desired item on a website.
- Every website must have a Contact Us section giving details of Facebook page, Twitter address, email address, phone number, official address, or any other link on the social media.

Conclusion

Google makes around 500 changes to its algorithm every year. Even a minor change in the algorithm can have a big impact on the performance of the website. Moreover, penalties for everything from obvious link spam to having too many ads on the website makes it mandatory for a company to go for a web audit semi-annually or annually.

Summary

- As the Internet is an insecure channel for exchanging private data or messages and intrusion or frauds like phishing are very common, some methods must be implemented to protect the data.
- Internet security ensures authenticated access of data that is exchanged over the Internet.
- Authorized or unauthorized users may modify the existing data, add wrong data, or delete some important data.

- The most common threats to data security comes either from use of malwares or through fraud like phishing.
- A worm, once installed, can connect to a remote computer over the Internet to download a more substantial piece of malicious software.
- Users must always use the latest operating system and antivirus software.
- In a man-in-the-middle (MIM) attack or network spoofing attack, the attacker intentionally inserts himself into a conversation between two persons.
- A good symmetric encryption algorithm is one that makes it very difficult, if not impossible, for attackers to decrypt the generated cipher text without knowing the key used for encryption.
- Ransomware is a subset of malware that locks the data (usually by encryption) on the victim's computer and asks for payment to decrypt that data and return access to the victim.
- Sniffing means monitoring and capturing data that is being transmitted over a network.
- Phishing is done to acquire sensitive information such as passwords, account numbers, and credit card details. In this technique, the fraudster constructs a fake website that looks similar to the legitimate website.
- Secure socket layer (SSL) is a security protocol that uses encryption of messages exchanged between a web server and a browser in an online communication.
- HTTPs ensures secure communication between a user's browser and a web server. A website supporting the HTTPs protocol has its URL beginning with https://.
- A website audit includes all activities to review the forms, information, content, graphics, evaluation of ranking, and the quality of a website. It is especially done to identify and improve any shortcomings that might have been otherwise ignored. During a website audit, factors like traffic are also monitored as it plays a key role in the success of a website and generation of sales.

Glossary

Accountability The activities of the authorized persons are documented to deter employees from wrongdoing

Antivirus A software that prevents, detects, and removes malicious software programs such as virus, worms, Trojan horses, spywares, adwares, and so on that are harmful to computer systems

Authentication Confirming the identity of the person requesting to gain access to a resource. This can be done by validating a username and password

Authorization Determining the set of actions that an authenticated person can perform on a resource. For example, an employee working in the Accounts department is not authorized to access the files of the Sales & Marketing department

Cookie Small-sized files that store information about an Internet user on her own computer

Cracker A person who breaks into the system by password cracking or by cracking the security measures implemented to protect the data

Cybercrime Any crime or illegal activity that involves a computer and a network

Cyberstalking Use of the Internet to stalk or harass an individual, an organization, or a specific group. Cyberstalking also includes monitoring someone's online activity

Cyberterrorism Disruptive use of IT by terrorists to attack networks, computer systems, and telecommunication infrastructures

Cyberwarfare Nations using IT to penetrate another nation's networks to cause damage or disruption

Cyberespionage The practice of using IT to obtain secret information without seeking permission from

Hacker A person who either breaks into the system for which they have no authorization or goes beyond their limits of legitimate access

Vulnerability assessment An audit or a check that is performed to identify potential vulnerabilities in a computer system or network

8. The generic term for all types of software designed with bad intentions to cause a threat to data security is _____.

9. The
that
soft
(a)
(b)
10. _____
(a)
(b)
11. Id
co
(a)
(b)
12. Id
sp
(a)
(b)
13. T
b
(
(
(
u
14. _____
(
r
15. _____
16. _____
17. _____
18. _____

- (a) virus (c) worm
(b) malware (d) Trojan horse
9. The generic term used for a piece of code that is usually embedded within a legitimate software is _____.
(a) virus (c) worm
(b) malware (d) Trojan horse
 10. _____ replicates itself.
(a) Virus (c) Adware
(b) Spyware (d) Trojan horse
 11. Identify the malware which can delete files or consume all the memory space.
(a) Virus (c) Adware
(b) Spyware (d) Trojan horse
 12. Identify the malware that needs a host file to spread itself.
(a) Virus (c) Adware
(b) Spyware (d) Trojan horse
 13. The _____ is loaded in memory from the boot record.
(a) MS Office
(b) operating system
(c) application software (d) utility software
 14. _____ viruses infect only executable files.
(a) Boot record (c) Stealth
(b) Program (d) Macro
 15. _____ viruses are a combination of boot viruses and program viruses.
(a) Multipartite (c) Stealth
(b) Polymorphic (d) Macro
 16. _____ virus will remove the virus code from an infected file when antivirus software is run.
(a) Multipartite (c) Stealth
(b) Polymorphic (d) Macro
 17. _____ viruses create copies during replication.
(a) Multipartite (c) Stealth
(b) Polymorphic (d) Macro
 18. _____ virus inserts itself in the computer's memory.
(a) Multipartite (c) Stealth
(b) Resident (d) Macro
 19. _____ virus is dormant when the file is not being executed.
(a) Direct action (c) Stealth
(b) Resident (d) Macro
 20. _____ virus changes the paths that indicate the location of a file.
(a) Overwrite (c) Network
(b) Directory (d) Cavity
 21. _____ virus multiplies through shared resources such as shared drives and files.
(a) Overwrite (c) Network
(b) Directory (d) Cavity
 22. When a file gets infected with a _____ virus, it seems missing or inaccessible to the users.
(a) overwrite (c) network
(b) directory (d) FAT
 23. _____ replicates itself.
(a) Worm (c) Adware
(b) Spyware (d) Trojan horse
 24. _____ is a malicious code that locates vulnerabilities on a computer to exploit them.
(a) Worm (c) Virus
(b) Spyware (d) Trojan horse
 25. _____ is a non-self-replicating malicious code that can even install a virus.
(a) Worm (c) Adware
(b) Spyware (d) Trojan Horse
 26. _____ may allow the computer to be accessed from a remote machine.
(a) Worm (c) Adware
(b) Spyware (d) Trojan horse
 27. _____ Trojan horse floods the target server with traffic to make it impossible for users to access certain websites.
(a) Denial-of-service attack
(b) File serving
(c) Keylogging
(d) Password stealing
 28. Antivirus can detect _____.
(a) viruses (c) ransomware
(b) Trojan horse (d) all of these
 29. _____ is usually used for tracking and storing a user's Internet browsing patterns.
(a) Worm (c) Virus

- (b) Spyware (d) Trojan horse
30. _____ does not transfer a user's personal information to another location.
(a) Worm (c) Adware
(b) Spyware (d) Trojan horse
31. _____ slows down a computer's speed.
(a) Virus (c) Adware
(b) Spyware (d) All of these
32. In _____, the attacker modifies the browser to permanently change the home page.
(a) denial-of-service
(b) spoofing
(c) browser hijacking
(d) sniffing
33. In _____, the attacker sends numerous external communication requests to the target machine.
(a) ping of flood (c) teardrop attack
(b) ping of death (d) mail bomb
34. A ping packet in the network that exceeds 65,535 bytes is sent in _____ attack.
(a) ping of flood (c) teardrop attack
(b) ping of death (d) mail bomb
35. In _____, the attacker puts a confusing sequence number in the packets.
(a) ping of flood (c) teardrop attack
(b) ping of death (d) mail bomb
36. In _____, unauthorized users send a large number of messages with large attachments to a particular server to fill its disk space.
(a) ping of flood (c) teardrop attack
(b) ping of death (d) mail bomb
37. In a _____ attack, the attacker inserts himself into a conversation between two people.
(a) man-in-the-middle (c) teardrop attack
(b) ping of death (d) mail bomb
38. In a _____ attack, the attacker impersonates the people in the conversation.
(a) ping of flood
(b) ping of death
(c) man-in-the-middle
(d) mail bomb
39. Man-in-the-middle attack can be prevented by _____.
(a) using an anti-spoofing software
(b) using HTTPs
(c) using packet filters
(d) all of these
40. In _____, the fraudster constructs a fake website that looks similar to the legitimate site and asks for the user's personal information to steal it.
(a) online auction fraud
(b) online purchase fraud
(c) phishing
(d) online intellectual property theft
41. Copy pasting content from a website without taking permissions from the author is _____.
(a) phishing
(b) violation of intellectual property
(c) work from home scam
(d) none of these
42. Link in a spam email may have _____.
(a) virus (c) bugs
(b) worm (d) all of these
43. A password cannot contain _____.
(a) characters
(b) digits
(c) special characters
(d) none of these
44. Which of the following is not necessary while working on the Internet?
(a) To go to a website, type its address in the address bar and not click on any link or cut and paste its address
(b) Delete cookies and history of web pages browsed
(c) Do not use an obsolete operating system
(d) None of these
45. Antivirus _____ to protect the computer from malware.
(a) deletes the file
(b) quarantines the file
(c) deletes the virus code
(d) all of these

46. Signa
files(a)
(b)
(c)
(d)47. In
dete
dict(a)
(b)48. The
the(a)
(b)49. W
an(a)
(b)

(c)

(c)

50. I
(a)

(c)

(c)

51.

52.

46. Signature-based antivirus software examines files when they are _____.
(a) created
(b) opened
(c) attached with an email
(d) all of these
47. In _____ type of antivirus, virus code is detected by searching for matches in a virus dictionary.
(a) signature-based (c) both of these
(b) heuristic based (d) none of these
48. The _____ type of antivirus protects the computer from brand-new viruses.
(a) signature-based (c) both of these
(b) heuristic based (d) none of these
49. Which of the following approaches does the antivirus not use?
(a) Emulating the beginning of the code of each executable
(b) Emulating the operating system and running the executable programs
(c) Working in the background and monitoring the computer's activity
(d) None of these
50. Identify the incorrect statement.
(a) Web-based antivirus scans should not be done as they are either spyware or Trojan horses.
(b) If there is more than one antivirus software installed, then only one of them should work in the interactive.
(c) Firewall exercises full control over data packets coming in and going out of the computer to the Internet.
(d) Anti-virus is configured with a set of rules that decide the packets to be accepted, which will be transmitted or received across the network.
51. Firewall can be a _____.
(a) hardware (c) both of these
(b) software (d) none of these
52. _____ type of firewall examines each data packet that either enters or leaves the network.
(a) Packet filtering
(b) Stateful
(c) Circuit-level gateway
(d) Proxy
53. _____ type of firewall records all connections passing through it.
(a) Packet filtering
(b) Stateful
(c) Circuit-level gateway
(d) Proxy
54. _____ firewall act as an intermediary for requests from one network to another.
(a) Packet filtering
(b) Stateful
(c) Circuit-level gateway
(d) Proxy
55. Firewalls can be implemented to block traffic for one or more _____.
(a) protocol (c) domain name
(b) IP address (d) all of these
56. _____ is the process of converting data into a cipher text.
(a) Encryption (c) Translation
(b) Decryption (d) Compression
57. Which property is to ensure that the sender cannot deny sending the message?
(a) Authentication (c) Integrity
(b) Non-repudiation (d) Consistency
58. Which property ensures that the message has not been modified or tampered during transmission?
(a) Authentication (c) Integrity
(b) Non-repudiation (d) Consistency
59. Which property verifies the originator of the message?
(a) Authentication (c) Integrity
(b) Non-repudiation (d) Consistency
60. _____ is the process of converting encrypted data back into its original form.
(a) Encryption (c) Translation
(b) Decryption (d) Compression
61. Which encryption algorithm uses a single key to encrypt messages?
(a) Symmetric (c) Both of these
(b) Asymmetric (d) None of these

62. Who knows the secret symmetric key?
 (a) Sender (c) Both of them
 (b) Receiver (d) None of them
63. Which of the following is not a disadvantage of the symmetric key algorithm?
 (a) Key may get into the hands of attackers.
 (b) No provision for authenticating the sender.
 (c) Data integrity cannot be assured.
 (d) The secret key can be a number, a word, or just a string of random letters.
64. In asymmetric key encryption algorithm, the original message is encrypted using _____ key.
 (a) public (c) symmetric
 (b) private (d) none of these
65. The encrypted message is decrypted with _____ key.
 (a) public (c) symmetric
 (b) private (d) none of these
66. Which of the following is not a feature of asymmetric encryption algorithm?
 (a) Both the sender and the receiver know about the secret key.
 (b) The two keys are mathematically related to each other.
 (c) The algorithm used for encryption is universally known.
 (d) It is impossible to compute the private key if the public key is known.
67. In a digital signature, the sender signs the message with his _____ key.
 (a) public (c) symmetric
 (b) private (d) none of these
68. In a digital envelope, the original message is encrypted with _____ key.
 (a) public (c) symmetric
 (b) private (d) none of these
69. In a digital envelope, symmetric key is encrypted with _____ key of the receiver.
 (a) public (c) symmetric
 (b) private (d) none of these
70. A digital certificate has the _____.
 (a) digital signature of the certificate-issuing authority
 (b) public key of the owner of the certificate
 (c) owner's identity
 (d) all of these
71. The job of certificate authority is to _____ certificates.
 (a) issue (c) both of these
 (b) revoke (d) none of these
72. _____ defines the degree to which data and other resources are accessible for use when required.
 (a) Reliability (c) Accessibility
 (b) Availability (d) Readability
73. Confirming the identity of the person requesting to gain access to a resource means _____.
 (a) authentication (c) reliability
 (b) authorization (d) accountability
74. _____ determines the set of actions that an authenticated person can perform on a resource.
 (a) Authentication (c) Reliability
 (b) Authorization (d) Accountability
75. Nations using IT to penetrate another nation's networks to cause damage or disruption, this is known as _____.
 (a) cyberterrorism (c) cybercrime
 (b) cyberwarfare (d) cyberespionage
76. The practice of using IT to obtain secret information without seeking permission from its owners is known as _____.
 (a) cyberterrorism (c) cybercrime
 (b) cyberwarfare (d) cyberespionage
77. _____ is a subset of malware which locks the data on the victim's computer and asks for payment to decrypt that data.
 (a) Ransomware (c) Spyware
 (b) Virus (d) Adware
78. Ransomware works by _____.
 (a) changing user's credentials
 (b) encrypting user's files and data
 (c) both of these
 (d) none of these
79. Payment in case of a ransomware attack is demanded in _____.
 (a) Dollars (c) Euro

80. Ransomware can get installed through _____.
- spam emails
 - infected apps
 - clicking pop-up messages
 - all of these
81. _____ is popular currency among cyber criminals because it is decentralized, unregulated, and practically impossible to trace.
- Dollars
 - Yen
 - Euro
 - Bitcoin
82. Which of the following is not a good way to control ransomware attacks?
- Downloading a good anti-virus software
 - Being extra cautious while clicking on links in emails
 - Installing a pop-up blocker on your computing device
 - Taking back up of computing devices on a regular basis
83. Which of the following is not correct?
- It is possible to delete a ransomware if the computer is not locked.
 - Anti-virus tools can detect and delete ransomware.
 - There are decryption tools that can be used in case of a ransomware attack.
 - Paying ransom guarantees that your files will be reusable.
84. _____ means tricking or deceiving computer systems or users.
- Spoofing
 - Sniffing
 - Hacking
 - Tracking
85. Spoofing is done by _____.
- hiding one's identity
 - faking the identity of another user on the Internet
 - DoS attack
 - both (a) and (b)
86. Identify the attack in which messages are either sent using a bogus e-mail address, IP address of a certain computer is masked, and a cyber criminal fakes an identity.
- Spoofing
 - Sniffing
 - Hacking
 - Tracking
87. Avoid trust relationship on the Internet to prevent _____.
- spoofing
 - sniffing
 - hacking
 - tracking
88. _____ means monitoring and capturing data that is being transmitted over a network.
- Spoofing
 - Sniffing
 - Hacking
 - Tracking
89. Legitimate use of sniffing does not involve _____.
- diagnosing network issues
 - troubleshooting network related issues
 - analysing network usage
 - extracting confidential data
90. Sniffing cannot be done by _____.
- internal user
 - external user
 - wireless user
 - none of these
91. In _____ sniffing, the sniffer just monitors and captures the data.
- active
 - passive
 - both of these
 - none of these
92. In _____ sniffing, the sniffer alters the data.
- active
 - passive
 - both of these
 - none of these
93. _____ sniffing is more difficult to detect and hence is more dangerous.
- Active
 - Passive
 - Both of these
 - None of these
94. Identify the protocol that protects computers and users from sniffing attack.
- HTTPs
 - SSL
 - TLS
 - All of these
95. Emails can be encrypted using _____ protocol which uses _____ encryption.
- PGP, symmetric
 - PGP, asymmetric
 - DES, symmetric
 - DES, asymmetric
96. _____ software can detect sniffing.
- Anti-sniffing
 - Antivirus
 - Anti-spyware
 - All of these
97. Identify the incorrect statement.

- (a) Computers that need not communicate directly with the Internet are placed in the DMZ.
 (b) Computers in DMZ are protected by firewalls.
 (c) A DMZ can also be set up on home networks.
 (d) DMZ can be a logical or a physical network.
98. SSL uses _____ of messages.
 (a) encryption (c) translation
 (b) decryption (d) compression
99. To create an SSL connection, one must have an SSL _____.
 (a) signature (c) degree
 (b) certificate (d) password
100. The SSL certificate contains the _____.
 (a) domain name
 (b) details of CA
 (c) applicant's address
 (d) all of these
101. The existence of an encrypted session is indicated by the presence of a _____.
 (a) lock icon in the address bar
 (b) green address bar
 (c) secured protocol
 (d) all of these
102. When browsing a website over an SSL connection, the URL begins with the _____ protocol.
 (a) HTTP (c) TLS
 (b) HTTPS (d) SFTP
103. SSL works with _____ encryption algorithm.
 (a) symmetric (c) both of these
 (b) asymmetric (d) none of these
104. Identify the incorrect statement.
 (a) SSL protects from phishing attack.
 (b) SEO gives higher ranking to websites using HTTPS.
 (c) Performance of a website degrades when it uses SSL.
 (d) HTTPS use SSL certificates.
105. HTTPS provide _____.
 (a) authentication (c) confidentiality
 (b) security (d) all of these
106. Identify the incorrect statement.
 (a) Without HTTPS, a government censor can choose to block certain pages of a site.
 (b) With HTTPS ISPs can no longer insert ads or inject code in the website.
 (c) On a website supporting HTTP, confidential data can be intercepted, spied on and even altered by anyone.
 (d) HTTPS exchanges data in an unencrypted form.
107. HTTPS encryption mechanism requires _____ key(s).
 (a) one (c) three
 (b) two (d) four
108. Identify the incorrect statement.
 (a) A website audit includes all activities to review the forms, information, content, graphics, evaluation of ranking, and the quality of a website.
 (b) During a website audit, traffic to a website is also monitored.
 (c) Website audit gives an insight into new graphic and strategic ideas.
 (d) Website audit does not reveal a website's ranking, bounce rates, and search percentages.
109. Website audit does not report _____.
 (a) response time (c) 404 errors
 (b) download time (d) none of these
110. To control duplication _____ is checked.
 (a) page title (c) keywords
 (b) page content (d) all of these
111. Identify the incorrect statement.
 (a) Web audit saves the website from penalization.
 (b) Web audit helps a website to discover malware infection.
 (c) Website audit should be an ongoing process that should be conducted at least once a year.

- (d) None of these
112. Users should not have to click more than _____ times to open a particular page.
- (a) one (c) three
(b) two (d) four
113. Identify the correct statement.
- (a) The homepage of a website should have detailed information about it.
(b) It is mandatory that Contact Us section be present in every website.
(c) The algorithm for ranking the websites is always the same.
(d) Every page should have a link, button, or any other way to return to the homepage.
114. Website audit should be done at least once _____.
- (a) in a week
(b) every fortnight
(c) in a month
(d) in six months
115. Cybercrime does not involve which of the following?
- (a) Using computer as a weapon to attack
(b) Using computer as an accessory to commit a crime
(c) Attacking a computer
(d) None of these
116. The most appropriate term used to denote the act of monitoring someone's online activity is _____.
- (a) cybercrime
(b) internet crime
(c) cyberstalking
(d) computer stalking
117. Hackers who take control over the system for personal gains are called _____ hackers.
- (a) red hat (c) grey hat
(b) white hat (d) black hat
118. _____ means gaining unauthorized access of a computer system.
- (a) Spoofing (c) Cracking
(b) Hacking (d) Attacking
119. _____ means analysing the data that is being transmitted over a network to gain access to some confidential information.
- (a) Spoofing (c) Cracking
(b) Hacking (d) Sniffing
120. _____ is done to acquire sensitive information such as passwords, account numbers, and credit card details by creating a fake website that looks similar to the legitimate website.
- (a) Phishing (c) Cracking
(b) Hacking (d) Sniffing
121. _____ file type is always safe to click and open.
- (a) .exe (c) docx
(b) .txt (d) .xlsx
122. _____ is the process of flooding the Internet with many copies of the same message for commercial advertising.
- (a) Phishing (c) Cracking
(b) Hacking (d) Spamming
123. Which attack targets the availability of web applications?
- (a) Phishing (c) DoS
(b) Sniffing (d) Spoofing
124. The 's' in https stands for _____.
- (a) safe (c) simple
(b) secure (d) straight
125. Website audit checks the _____.
- (a) content (c) performance
(b) graphics (d) all of these
126. Which of the following is correct?
- (a) The home page should have maximum information.
(b) Cookies are small-sized files that store information about an Internet user on his/her own computer.
(c) Password cracker is an ethical hacker.
(d) Authorization means confirming the identity of the person requesting to gain access to a resource.

127. The degree to which the accuracy and completeness of information and computer software are protected against unauthorized access ensures _____

- (a) availability (c) integrity
(b) confidentiality (d) accountability

128. _____ means monitoring someone's online activity.

- (a) Cybercrime (c) Cyberwarfare
(b) Cyberstalking (d) Cyberespionage

Answers to Multiple-choice Questions

- | | | | | | | | | | |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1. (d) | 2. (d) | 3. (d) | 4. (d) | 5. (d) | 6. (d) | 7. (d) | 8. (b) | 9. (b) | 10. (a) |
| 11. (a) | 12. (a) | 13. (b) | 14. (b) | 15. (a) | 16. (c) | 17. (b) | 18. (b) | 19. (a) | 20. (b) |
| 21. (c) | 22. (d) | 23. (a) | 24. (a) | 25. (d) | 26. (d) | 27. (a) | 28. (d) | 29. (b) | 30. (c) |
| 31. (d) | 32. (c) | 33. (a) | 34. (b) | 35. (c) | 36. (d) | 37. (a) | 38. (c) | 39. (d) | 40. (c) |
| 41. (b) | 42. (d) | 43. (d) | 44. (d) | 45. (d) | 46. (d) | 47. (a) | 48. (b) | 49. (d) | 50. (d) |
| 51. (c) | 52. (a) | 53. (b) | 54. (d) | 55. (d) | 56. (a) | 57. (b) | 58. (c) | 59. (a) | 60. (b) |
| 61. (a) | 62. (c) | 63. (d) | 64. (a) | 65. (b) | 66. (a) | 67. (b) | 68. (c) | 69. (a) | 70. (d) |
| 71. (c) | 72. (b) | 73. (a) | 74. (b) | 75. (b) | 76. (d) | 77. (a) | 78. (c) | 79. (d) | 80. (d) |
| 81. (d) | 82. (d) | 83. (d) | 84. (a) | 85. (d) | 86. (a) | 87. (a) | 88. (b) | 89. (d) | 90. (d) |
| 91. (b) | 92. (a) | 93. (b) | 94. (d) | 95. (b) | 96. (d) | 97. (a) | 98. (a) | 99. (b) | 100. (d) |
| 101. (d) | 102. (b) | 103. (b) | 104. (c) | 105. (d) | 106. (d) | 107. (b) | 108. (d) | 109. (d) | 110. (d) |
| 111. (d) | 112. (c) | 113. (d) | 114. (d) | 115. (d) | 116. (c) | 117. (d) | 118. (b) | 119. (d) | 120. (a) |
| 121. (b) | 122. (d) | 123. (c) | 124. (b) | 125. (d) | 126. (b) | 127. (c) | 128. (b) | | |

IT Act 2000 and cybercrimes

CHAPTER

5

Syllabus Mapping

IT Act 2000—Definitions of different terms, digital signature, electronic governance, attribution, acknowledgement and dispatch of electronic records, regulation of certifying authorities, digital signature certificates, duties of subscribers, penalties and adjudication, appellate tribunal, offences and cybercrimes.

Unit

Module I
Unit 5

5.1 INFORMATION TECHNOLOGY ACT 2000

High-speed Internet connectivity has no doubt brought about a communication revolution. However, on the flipside, it has led to an increase in online crimes. A rise in these offences necessitated effective laws for protection online. Keeping this in mind, the Indian Parliament passed the Information Technology Act 2000 (No. 21 of 2000). This was conceptualized on the United Nations Commissions on International Trade Law (UNCITRAL) model.

The Government of India enacted the Information Technology (IT) Act with the objective to deliver and facilitate lawful electronic, digital, and online transactions, and mitigate cybercrimes. It defines the offences, along with the penalties for each category of offence, in detail.

The IT Act of India provides legal recognition for transactions involving exchange of electronic data and other means of electronic communication, commonly referred to as electronic commerce (or e-commerce). E-commerce involves alternatives to paper-based methods of communication and storage of information.

5.2 SALIENT FEATURES OF IT ACT

Some of the key features of the IT Act are listed here:

1. Digital signatures have been replaced with electronic signatures.
2. A detailed note on offences, penalties, and breaches is given.
3. It talks about the justice dispensation systems for cybercrimes.
4. It provides details for the constitution of the Cyber Regulations Advisory Committee.
5. The Act is based on The Indian Penal Code 1860, The Indian Evidence Act 1872, The Bankers' Books Evidence Act 1891, The Reserve Bank of India Act 1934, etc.
6. It adds a provision to Section 81, which states that *nothing contained in the Act shall restrict any person from exercising any right conferred under the Copyright Act 1957.*

5.3 DEFINITION

The term 'Act' here means the Information Technology Act 2000. The IT Act extends to the whole of India and is also applicable to any offence or contravention committed outside India by any person.

The Act also clarifies that no clause given in this Act shall apply to

1. a negotiable instrument as defined in Section 13 of the Negotiable Instruments Act 1881,
2. a power-of-attorney as defined in Section 1A of the Powers-of-Attorney Act 1882,
3. a trust as defined in Section 3 of the Indian Trusts Act 1882,
4. a will as defined in Clause (h) of Section 2 of the Indian Succession Act 1925 including any other testamentary disposition by whatever name called,
5. any contract for the sale or conveyance of immovable property or any interest in such property, and
6. any such class of documents or transactions as may be notified by the Central government in the official gazette.

To read and understand the IT Act, one must know the context in which certain terms are used. The following are the key terms.

Act The Information Technology Act 2000 (21 of 2000)

Applicant The certifying authority

Auditor Any internationally accredited computer security professional or agency appointed by the certifying authority and recognized by the controller for doing a technical audit of the work done by the certifying authority

Controller The controller of certifying authorities appointed under Sub-section (1) of Section 17 of the Act

Digital signature certificate The digital signature certificate issued under Sub-section (4) of Section 35 of the Act

Information asset All information resources utilized in the organization for conducting its business, including all the information, applications (i.e., software developed or purchased), and technology (hardware, system software, and networks)

Licence A licence (or permission) granted to certifying authorities for the issue of digital signature certificates under the defined rules

Licensed certifying authority The certifying authority who has the licence to issue digital signature certificates

Person An individual, a company, an association, a body of individuals, or the Central government or state government or any of the ministries or departments, agencies or authorities of such governments

Schedule The Act has two schedules, the First Schedule dealing with documents or transactions to which the Act shall not apply, and the Second Schedule, dealing with electronic signatures or electronic authentication techniques and procedures

Subscriber identity verification method The method used to verify and authenticate the identity of a subscriber

Trusted person Any person who has

1. direct responsibilities towards the day-to-day operations, security, and performance of those business activities that are regulated under the Act or these rules in respect of a certifying authority; or

2. duties directly involving the issuance, renewal, suspension, revocation of digital signature certificates (including the identification of any person requesting a digital signature certificate from a licensed certifying authority), creation of private keys, or administration of a certifying authority's computing facilities.

Cyber café Any facility from where Internet access is offered by any person in the ordinary course of business to the members of the public



The IT Act contains 13 chapters and 90 sections.

5.3.1 Amendments to IT Act

The IT Act has undergone amendments, which have been provided in Schedule 1-4.

1. The First Schedule has widened the scope of the term *document* to bring within its ambit electronic documents.
2. The Second Schedule pertains to the inclusion of electronic documents in the definition of evidence.
3. The Third Schedule amends the Bankers' Books Evidence Act. As per this amendment, a *banker's book* now includes printouts of data stored in a floppy, disc, tape, or any other form of electromagnetic data storage device.
4. The Fourth Schedule amends the Reserve Bank of India Act. It regulates the transfer of funds through electronic means between the banks or between the banks and other financial institutions.

5.4 DIGITAL SIGNATURE

We use a signature to validate a document. Similarly, a digital signature is a technique that is used to validate the legitimacy of a digital message or a document. A valid digital signature assures the recipient that the message was generated by a known sender. It also rules out the possibility of non-repudiation (the sender cannot deny having sent the message).

Electronic signature An electronic signature or e-signature also plays the same role. It ensures that the person who has signed the document has created the message. It also indicates that the person accepts the purposes recorded in the document.



In many companies, digital seals are also required for another layer of authentication and security.

Digital signature to electronic signature The term digital signature was defined in the old IT Act 2000. However, the term *electronic signature* was defined by the amended Act (IT Act 2008). The concept of electronic signature is broader than digital signature. Section 3 of the Act states that the verification of electronic records by electronic signatures or electronic authentication techniques shall be considered reliable.

According to UNCITRAL, the following technologies are presently in use—digital signature, biometric device, PIN, password, scanned handwritten signature, signature by digital pen, and clickable 'OK', 'I Accept', or 'I Agree' click boxes.

5.5 ELECTRONIC GOVERNANCE

Electronic governance or e-governance is the application of information and communication technology (ICT), especially for delivering government services or exchanging information between government-to-citizen (G2C), government-to-business (G2B), government-to-government (G2G), government-to-employees (G2E) as well as back office processes. E-governance enables citizens to use government services in a convenient, efficient, and transparent manner. To support the e-governance system, IT Act has the following provisions.

Legal recognition of digital signatures According to the law, information or any other matter shall be authenticated by affixing a signature. Documents shall be signed or bear the signature of any person. Such a requirement will be satisfied, if such information or matter is authenticated by means of a digital signature affixed in such manner as may be prescribed by the Central government.

The appropriate government (state or central) may define rules regarding the manner and format in which such electronic records shall be filed, created, or issued. Rules may also define the manner or method of payment of any fee or charges for filing, creation, or issue of any electronic record under Clause (a).

1. the information contained therein remains accessible so as to be usable for a subsequent reference;
2. the electronic record is retained in the format in which it was originally generated, sent, or received or in a format which can be demonstrated to represent accurately the information originally generated, sent, or received; and
3. the details which provide for the identification of the origin, destination, date, and time of dispatch or receipt of such electronic records that are available in the electronic record.

Publication of rule, regulation, etc., in electronic gazette The Act also states that any rule, regulation, order, bye-law, notification or any other matter shall be published in the official gazette or electronic gazette provided that the date of publication shall be deemed to be the date of the gazette which was first published in any form.

Not to confer right to insist document should be accepted in electronic form No clause stated above shall confer (or grant) a right upon any person to insist that any ministry or department of the Central government or the state government, or any authority or body established by or under any law or controlled or funded by the central or state government should accept, issue, create, retain, and preserve any document in the form of electronic records or effect any monetary transaction in the electronic form.

Power to make rules by Central government in respect of digital signature The Central government may prescribe

1. the type of digital signature;
2. format and the way in which the digital signature shall be affixed;
3. the procedure which enables identification of the person affixing the digital signature;
4. control processes and procedures to ensure adequate integrity, security, and confidentiality of electronic records or payments; and
5. any other issue which is used to give legal effect to digital signatures.

5.6 ATTRIBUTION, ACKNOWLEDGMENT, AND DISPATCH OF ELECTRONIC RECORDS

This section mainly deals with electronic contracts. Attribution means the act of establishing a particular person as the creator of a work. With respect to the IT Act, attribution of electronic records means fixing identity of sender and receiver. It states that originator is a person who sends or generates any electronic record and receiver means the addressee. For example if 'A' sends an email to 'B', then 'A' is the sender or originator and 'B' is the addressee. In case of postal communication or paper communication, it is easy to identify originator and addressee but in electronic communication it is not that straight forward because the electronic record can be sent by the originator himself, or by the person who has been authorized by the originator, or by an information system that the originator has authenticated.

Attribution of electronic records An electronic record shall be attributed to the originator in the following cases:

1. If the record was sent by the originator himself.
2. If the record was sent by a person who was authorized to act on and send the electronic record on behalf of the originator.
3. If the record was automatically sent by an information system programmed by or on behalf of the originator.

Acknowledgment of receipt If a particular form and method of the receipt of electronic record has not been mutually decided upon by the originator and addressee, an acknowledgment may be given by

1. any communication by the addressee, automated or otherwise; or
2. in any other way sufficient to indicate to the originator that the electronic record has been received.

If the originator has clearly stated that the electronic record shall be binding only on receipt of an acknowledgment of the electronic record, unless acknowledgment has been received by the originator, the electronic record shall be deemed to have been never sent by the originator.

If the originator has not stated anything regarding the binding of electronic record only on receipt of acknowledgment, if the acknowledgment has not been received by the originator within the time specified or agreed, or if no time has been specified or agreed to within a reasonable time, the originator may give a notice to the addressee stating that no acknowledgment has been received and also specifying a reasonable time by which the acknowledgment must be received. Even after doing so, if no acknowledgment is received within the aforesaid time limit, the originator may, after giving a notice to the addressee, treat the electronic record as though it has never been sent.

Time and place of dispatch and receipt of electronic record If the time of dispatch has not been agreed upon between the originator and the addressee, the dispatch of an electronic record occurs when it enters a computer resource outside the control of the originator. If the meaning of receipt or dispatch has not been mutually discussed and agreed upon between the originator and the addressee, two scenarios can occur.

1. First, if the addressee has designated a computer resource for the purpose of receiving electronic records, the receipt of the electronic record occurs at the time the record enters the designated computer resource.
2. Second, if the addressee has not designated a computer resource, the receipt occurs at the time the electronic record is retrieved by the addressee.

If the place of dispatch has not been mutually discussed and agreed upon between the originator and the addressee, an electronic record is deemed to be dispatched at the place where the originator has his/her place of business, and is deemed to have been received at the place where the addressee has his/her place of business.

If the place where the computer resource is located is different from the place the electronic record is deemed to have been received, and if the originator or the addressee has more than one place of business, the principal place of business shall be the place of business.

However, if the originator or the addressee does not have a place of business, the place of residence (or a place where the business is registered) shall be deemed to be the place of business.

5.7 REGULATION OF CERTIFYING AUTHORITIES

The working of the computer, computer network, and computer system must have an identification strategy to ascertain the integrity, confidentiality, and authentication of electronic communication. Usually, the electronic environment uses digital signatures to identify and prove transactions. Therefore, there is a need for trusted third parties to authenticate that a digital signature belongs to a specific signer and distribute them the public keys. So, important terms in this section include: the *Certifying Authority* which issues Digital Signature Certificates by authenticating the subscriber's identity. However, it can issue certificates only after obtaining a licence from the *Controller of Certifying Authorities* or *Root Certifying Authority of India* (RCAI).

Appointment of Controller and Other Officers

The Central government may, by notification in the official gazette, appoint a controller of certifying authorities and a number of deputy controllers and assistant controllers as it deems fit for the purposes of this Act.

1. The controller will perform all his/her functions under the general control and directions of the Central government.
2. The controller will assign tasks to the deputy controllers and assistant controllers. They will therefore work under the general superintendence and control of the controller.
3. The qualifications, experience, and terms and conditions of service of the controller, deputy controllers, and assistant controllers shall be laid out by the Central government.

The head office and branch office of the office of the controller shall be determined by the Central government. There shall be a seal of the office of the controller.

Functions of Controller

The controller may perform all or any of the following functions:

1. Supervise the activities of the certifying authorities.
2. Certify public keys of the certifying authorities.
3. Lay down the standards to be maintained by the certifying authorities. Specify the qualifications and experience that should be possessed by the employees of the certifying authorities.
4. Specify the conditions under which the certifying authorities shall conduct their business.
5. Specify the contents of written, printed, or visual materials and advertisements that may be distributed or used in respect of a digital signature certificate and the public key.
6. Lay out the format and content of a digital signature certificate and the key.
7. Specify the manner in which accounts shall be maintained by the certifying authorities.
8. Specify the terms and conditions under which auditors may be appointed and the remuneration to be paid to them.
9. Facilitate the establishment of any electronic system by a certifying authority either solely or jointly with other certifying authorities and regulation of such systems.
10. Specify the manner in which the certifying authorities will deal with the subscribers.
11. Resolve any conflict of interests between the certifying authorities and the subscribers.
12. List the duties of the certifying authorities.
13. Maintain a database of the disclosure record of every certifying authority consisting of all the particulars specified by regulations. This database should be made accessible to the public.

5.8 DIGITAL SIGNATURE CERTIFICATES

We know that a digital signature certificate is issued by the certifying authorities for validating and certifying the identity of the person holding it. The certificate contains information about the user's name, pin code, country, email address, date of issuance of certificate, and name of the certifying authority.

Certifying Authority to Issue Digital Signature Certificate

Any person may apply for the issue of a digital signature by the certifying authority in a form as prescribed by the Central government.

Every application shall be accompanied by the specified fee (not exceeding twenty-five thousand rupees) as prescribed by the Central government. This fee will be paid to the certifying authority.

Every application shall be accompanied by a certification practice statement. In the absence of such a statement, a statement containing particulars as specified by the regulations should be attached.

On receipt of an application, the certifying authority may, after consideration of the certification practice statement or the other statement, make enquiries to validate the statement. And in case it deems fit, the certifying authority grants the digital signature certificate. Otherwise it rejects the application.

To determine whether to grant the digital signature or reject the application, the certifying authority first ensures the following:

1. Whether the applicant holds the private key corresponding to the public key to be listed in the digital signature certificate
2. Whether the applicant holds a private key, which is capable of creating a digital signature

3. Whether the public key to be listed in the certificate can verify a digital signature affixed by the private key held by the applicant

Moreover, an application cannot be rejected unless the applicant has been given reasonable opportunity to show cause against the proposed rejection.

Representations Upon Issuance of Digital Signature Certificate

While issuing a digital certificate, the certifying authority shall certify that

1. it is being provided as per the rules and regulations specified in this Act;
2. the digital signature certificate is published, is available to the person relying on it, and has been accepted by the subscriber;
3. the subscriber holds the private key corresponding to the public key which is listed in the digital signature certificate;
4. the subscriber's public key and private key constitute a functioning key pair;
5. the information given in the digital signature certificate is accurate; and
6. it has no knowledge of any fact which, had it been included in the digital signature certificate, would adversely affect the reliability of the representations made in the clauses specified herein.

Suspension of Digital Signature Certificate

The certifying authority which issued a digital signature certificate may suspend it in the following scenarios:

1. If it receives a request either from the subscriber listed in the digital signature certificate, or any person duly authorized to act on behalf of that subscriber.
2. If the digital signature certificate should be suspended in public interest.

However, the digital signature certificate shall not be suspended for a period exceeding 15 days unless the subscriber has been given the opportunity of being heard in the matter.

Further, while suspending the digital signature certificate, the certifying authority shall communicate the same to the subscriber.

Revocation of Digital Signature Certificate

A certifying authority may revoke a digital signature certificate issued by it under the following circumstances:

1. If either the subscriber or any other person authorized by him/her makes a request to revoke it
2. Upon the death of the subscriber
3. Upon the dissolution of the firm or winding-up of the company where the subscriber is a firm or a company
4. If the certifying authority finds that a material fact represented in the digital signature certificate is false or has been concealed
5. If the requirement for issuance of the digital signature certificate was not satisfied
6. If the certifying authority's private key or security system has been compromised in a way that can affect the reliability of the digital signature certificate

However, a digital signature certificate shall not be revoked unless the subscriber has been given the opportunity of being heard in the matter. Moreover, while revoking the digital signature certificate, the certifying authority shall communicate the same to the subscriber.

Notice of Suspension or Revocation

Whenever a digital signature certificate is suspended or revoked, the certifying authority shall publish a notice of such suspension or revocation.

5.9 DUTIES OF SUBSCRIBERS

In general, a subscriber is the person who pays certain subscription amount to avail some kind of service. Similarly, in the Public Key Infrastructure (PKI), subscriber is the customer who pays to become a member of the Digital Signature Certificate 'club'. After subscribing for the Digital Signature Certificate, a subscriber is bound to fulfil the following responsibilities.

Generation of Key Pair

When a digital signature certificate is issued, the public key of the subscriber is mentioned in the certificate. Now, it shall be the duty of the subscriber to generate his/her private key by applying the security procedure. The key pair is generated in such a way that any text encrypted using a public key can be decrypted by its corresponding private key.

Acceptance of Digital Signature Certificate

1. A subscriber shall be deemed to have accepted a digital signature certificate if he/she publishes or authorizes the publication of a digital signature certificate to one or more persons or in a repository, or approves the digital signature certificate in any manner.
2. By accepting a digital signature, the subscriber certifies to all those who rely on the information contained in the digital signature certificate that
 - (a) the subscriber has the private key corresponding to the public key specified in the digital signature certificate;
 - (b) all representations made by the subscriber to the certifying authority and all the information contained in the digital signature certificate is true; and
 - (c) all the information in the digital signature certificate that is within the knowledge of the subscriber is true.

Control of Private Key

1. Every subscriber shall take care to retain control of the private key corresponding to the public key listed in his/her digital signature certificate. The subscriber shall take all steps necessary to prevent the disclosure of the private key to a person not authorized to use it in any way.
2. If the private key corresponding to the public key listed in the digital signature certificate has been compromised, then, without any delay, the subscriber shall communicate the same to the certifying authority in the way as specified by the regulations.

The clause also states that the subscriber shall be considered liable until he/she informs the certifying authority that the private key has been compromised.

5.10 PENALTIES AND ADJUDICATION

Penalty for Damage to Computer, and Computer System

If any person without permission of the owner or any other person who is in charge of a computer, computer system, or computer network accesses it, downloads files, copies, or extracts any data, introduces or causes introduction of a computer virus in it, damages or causes it to be damaged in any way, disrupts or causes its disruption, denies or causes the denial of access to any person authorized to access it is liable to pay a penalty.

The person who provides any assistance to any other person to facilitate access to a computer, computer system, or computer network or charges the services availed of by a person to the account of another person by tampering with or manipulating any computer, computer system, or computer network shall be liable to pay damages by way of compensation not exceeding one crore rupees to the person so affected.

Definitions in This Clause

Computer contaminant It refers to a set of computer instructions that are designed

1. to modify, destroy, record, or transmit data or programs residing within a computer, computer system, or computer network; or
2. by any means to usurp (take illegally or by force) the normal operation of the computer, computer system, or computer network.

Computer database It refers to a representation of information, knowledge, facts, concepts, or instructions in text, image, audio, and video that are being prepared or have been prepared in a formalized manner or have been produced by a computer, computer system, or computer network and are intended for use in a computer, computer system, or computer network.

Computer virus It refers to any computer instruction, information, data, or program that destroys, damages, adversely degrades, affects the performance of computer resources, or attaches itself to another computer resource and operates when a program, data, or instruction is executed or some other event takes place in that computer resource.

Damage It refers to the destruction, alteration, deletion, addition, modification, or re-arrangement of any computer resource by any means.

Penalty for Failure to Furnish Information and Returns

If any person who is required to furnish any document, return, or report to the controller or the certifying authority fails to do the same, he/she shall be liable to a penalty not exceeding one lakh and fifty thousand rupees for each such failure.

If a person who is supposed to file any return or furnish any information, books, or other documents within the time fails to do so, he/she shall be liable to a penalty not exceeding five thousand rupees for every day during which such failure continues.

If a person who is expected to maintain books of accounts or records fails to do the same, he/she shall be liable to a penalty not exceeding ten thousand rupees for every day during which the failure continues.

Residuary Penalty

Whoever contravenes any rules or regulations made under this Act, for the contravention of which no penalty has been separately provided, shall be liable to pay a compensation not exceeding twenty-five thousand rupees to the person affected by such contravention or a penalty not exceeding twenty-five thousand rupees.

Power to Adjudicate

For the purpose of adjudging (make a formal judgment on a disputed matter) if any person has committed a contravention (an action which goes against a law, treaty, or other ruling) of any of the provisions of this Act or of any rule, regulation, direction, or order made under the Central government shall appoint any officer not below the rank of a director to the Government of India or an equivalent officer of a state government to be an adjudicating officer for holding an inquiry in the manner prescribed by the Central government.

The adjudicating officer shall give the person reasonable opportunity for making a representation in the matter. If, after an inquiry, the officer finds that the person has committed the contravention, he/she may impose such penalty, award, or compensation as he/she thinks fit in accordance with the provisions of that section.

No person shall be appointed as an adjudicating officer unless he/she is experienced in the field of information technology and legal matters as prescribed by the Central government.

If more than one adjudicating officers are appointed, the Central government shall specify by order the matters and places with respect to which such officers shall exercise their jurisdiction.

Every adjudicating officer shall have the powers of a civil court which are conferred on the Cyber Appellate Tribunal.

Factors to be Taken into Account by Adjudicating Officer

While adjudging the compensation, the adjudicating officer shall consider the following factors:

1. the amount of gain of unfair advantage, whenever quantifiable, made as a result of the default;
2. the amount of loss caused to any person as a result of the default; and
3. the repetitive nature of the default.

5.11 CYBER REGULATIONS APPELLATE TRIBUNAL

Tribunals are an important feature of the Indian judicial system as they resolve disputes especially those related to technical areas. For this, they require technical knowledge and familiarity with specialized factual scenarios.

5.11.1 Establishment

The Central government shall, by notification, establish one or more appellate tribunals to be known as the Cyber Regulations Appellate Tribunal. The Central government shall also specify in the notification the matters and places in relation to which the tribunal may exercise jurisdiction.

5.11.2 Composition

The Cyber Appellate Tribunal shall consist of only one person (to be known as the Presiding Officer) to be appointed, by notification, by the Central government.

5.11.3 Qualifications for Appointment as Presiding Officer of Cyber Appellate Tribunal

A person shall be qualified for appointment as the Presiding Officer of Cyber Appellate Tribunal if he/she is, has been, or is qualified to be a judge of a high court; or if he/she is, or has been, a member of the Indian Legal Service and is holding or has held a post in Grade I of that service for at least three years.

5.11.4 Term of Office

The Presiding Officer of Cyber Appellate Tribunal shall hold office for a term of five years from the date on which he/she enters upon his/her office or until he/she attains the age of sixty-five years, whichever is earlier.

5.11.5 Salary, Allowance, and Other Terms of Service of Presiding Officer

The salary and allowances payable to, and the other terms and conditions of service including pension, gratuity, and other retirement benefits of the Presiding Officer of Cyber Appellate Tribunal shall not be varied to his/her disadvantage after appointment.

5.11.6 Filling-up of Vacancies

If for any reason other than temporary absence, a vacancy occurs in the office of the Presiding Officer of a Cyber Appellate Tribunal, then the Central government shall appoint another person and the proceedings may be continued before the Cyber Appellate Tribunal from the state at which the vacancy is filled.

5.11.7 Resignation and Removal

The Presiding Officer of Cyber Appellate Tribunal may, by notice in writing under his/her hand addressed to the Central government, resign his/her office. However, the Central government can permit him/her to relinquish his/her office sooner, but make him/her continue to hold office until the expiry of three months from the date of receipt of such notice or until a person duly appointed as his/her successor enters upon his/her office or until the expiry of his/her term of office, whichever is the earliest.

The Presiding Officer shall not be removed from his/her office except by an order by the Central government on the ground of proven misbehaviour or incapacity after an inquiry made by a judge of the Supreme Court. The Presiding Officer shall be informed of the charges against him/her and given reasonable opportunity of being heard in respect of those charges.

The Central government may, by rules, regulate the procedure for the investigation of misbehavior or incapacity of the Presiding Officer.

5.11.8 Orders Constituting Appellate Tribunal to be Final and not Invalidate Proceedings

The order of the Central government appointing any person as the Presiding Officer of a Cyber Appellate Tribunal shall not be questioned in any manner. Moreover, no act or proceeding before a Cyber Appellate

Tribunal shall
tion of the Cyb

5.11.9 Sta

The Central g
officers and en
Officer. Their
government.

5.11.10 A

Any person a
Cyber Appel
Appellate Tri

The appea
by the contro
the prescribe
after 45 days
cause for no

On receip
heard. It the
against.

The Cyb
the concern

The app
and endea
appeal.

5.11.11

The Cyber
1908, but s
Appellate
have its sit

To perform
Code of C

1. Summ
2. Findi
3. Rece
4. Issui
5. Revi
6. Disn
7. Any

Every pr

Tribunal shall be called 'questionable' in any manner merely on the basis of any defect in the constitution of the Cyber Appellate Tribunal.

5.11.9 Staff of Cyber Appellate Tribunal

The Central government shall provide the Cyber Appellate Tribunal with officers and employees. The officers and employees shall discharge their functions under the general superintendence of the Presiding Officer. Their salaries, any allowances, and other conditions of service shall be prescribed by the Central government.

5.11.10 Appeal to Cyber Regulations Appellate Tribunal

Any person aggrieved by an order made by the controller or an adjudicating officer may appeal to a Cyber Appellate Tribunal having jurisdiction in the matter. However, no appeal shall lie to the Cyber Appellate Tribunal from an order made by an adjudicating officer with the consent of the parties.

The appeal shall be filed within a period of 45 days from the date on which a copy of the order made by the controller or the adjudicating officer is received by the person aggrieved. The appeal shall be in the prescribed format and accompanied by the specified fee. In case the person aggrieved files the appeal after 45 days, it will be accepted only if the Cyber Appellate Tribunal is satisfied that there was sufficient cause for not filing it within that period.

On receipt of an appeal, the Cyber Appellate Tribunal gives both the parties an opportunity of being heard. It then passes orders as it thinks fit, confirming, modifying, or setting aside the order appealed against.

The Cyber Appellate Tribunal shall then send a copy of every order made by it to the parties and to the concerned controller or adjudicating officer.

The appeal filed before the Cyber Appellate Tribunal shall be dealt with as expeditiously as possible and endeavour shall be made to dispose off the appeal within six months from the date of receipt of the appeal.

5.11.11 Procedure and Powers

The Cyber Appellate Tribunal shall not be bound by the procedure given by the Code of Civil Procedure 1908, but shall be guided by the principles of natural justice and other provisions of this Act. The Cyber Appellate Tribunal shall have powers to regulate its own procedure including the place at which it shall have its sittings.

To perform its functions, the tribunal shall have the same powers as are vested in a civil court under the Code of Civil Procedure 1908 while trying a suit, in context of the following matters:

1. Summoning and enforcing the attendance of any person and examining him/her on oath
2. Finding and producing documents or other electronic records
3. Receiving evidence on affidavits
4. Issuing commissions for the examination of witnesses or documents
5. Reviewing its decisions
6. Dismissing an application for default or deciding it ex parte
7. Any other matter which may be prescribed

Every proceeding before the Cyber Appellate Tribunal shall be deemed to be a judicial proceeding.

5.11.12 Right to Legal Representation

The appellant may either appear in person or authorize one or more legal practitioners or any of its officers to present his/her case before the Cyber Appellate Tribunal.

5.11.13 Civil Court Not to Have Jurisdiction

No court shall have jurisdiction to entertain any suit or proceeding of any matter which an adjudicating officer or the Cyber Appellate Tribunal constituted under this Act is empowered by. No court shall give any order in respect of any action taken or to be taken by the adjudicating officer or the Cyber Appellate Tribunal according to any power conferred by or under this Act.

5.11.14 Appeal to High Court

If a person is aggrieved by any decision or order of the Cyber Appellate Tribunal, he/she may file an appeal to the High Court within 60 days from the date of communication of the decision or the order to him/her. If the aggrieved person fails to appeal within 60 days, his/her appeal will be accepted only if the High Court is satisfied that the appellant was prevented by sufficient cause from filing the appeal within the said period. In that case, the applicant is allowed to file the appeal within a further period not exceeding 60 days.

5.11.15 Compounding of Contraventions

Any contravention (offence under this law) may be compounded by the controller or other officer specially authorized by him/her or by the adjudicating officer. However, the sum shall not, in any case, exceed the maximum amount of the penalty which may be imposed under this Act for the contravention so compounded.

Moreover, nothing in the sub-section shall apply to a person who commits the same or similar contravention within a period of three years from the date on which the first contravention committed by him was compounded.

Any second or subsequent contravention committed after the expiry of a period of three years from the date on which the contravention was previously compounded, shall be deemed to be the first contravention.

5.11.16 Recovery of Penalty

A penalty imposed under this Act, if not paid, shall be recovered as an arrear of land revenue and the licence, the digital signature certificate, and the case shall be suspended till the penalty is paid.

5.12 OFFENCES AND CYBERCRIMES

These days, criminals take advantage of IT to affect society in different ways. In Chapter 4, we have discussed cybercrimes in detail. In this section, we will recapitulate the concepts to learn what steps have been taken by the Government of India to protect its citizens against such attacks.

The term *cybercrime* covers a wide range of criminal activities that makes use of a computer. Some common examples of such crimes include identity theft, financial fraud, website defacements, cyber bullying, hacking of databases, theft of intellectual property, cyber harassment (harassing a person based

on sexual,
unauthoriz
Reasons f
ments and
Effects of
property,
Avoiding
and action
Moreover
ily mean

5.12.1

A perso
for a co
require
or fine
Her
layout,

5.12.2

A per
who c
or uti
onme
punis

5.12.3

Any
who
anne
ill-v
a te

5.12.4

A p
kn
sh
ma

on sexual, racial, religious, or other factors), software piracy, transmission of harmful programs, and unauthorized access of computerized information to name a few.

Reasons for cybercrime Mischief, to become popular, financial gain, to take revenge, forging documents and messages for personal gain.

Effects of cybercrime A single, successful cyberattack can result in financial losses, theft of intellectual property, and loss of consumer confidence and trust.

Avoiding cybercrime To avoid cyberattacks, every user must be trained and made aware of cyber threats and actions that can be taken to protect their information.

Moreover, the Government of India introduced The Information Technology Act 2000, which is primarily meant to be a legislation to promote e-commerce.

5.12.1 Tampering with Computer Source Documents

A person who knowingly or intentionally conceals, destroys, or alters any computer source code used for a computer, computer program, computer system, or computer network, when the source code is required to be kept or maintained by law for the time being in force, shall be imprisoned for three years, or fined with an amount which may extend up to two lakh rupees, or be punished with both.

Here, the term *computer source code* means the listing of programs and commands, and the design, layout, and program analysis of computer resources in any form.

5.12.2 Hacking with Computer System

A person who knowingly or intentionally causes wrongful loss or damage to the public or any person who destroys, deletes, or alters any information residing in a computer resource or diminishes its value or utility, or permanently damages it by any means, commits hacking and shall be punished with imprisonment up to three years, or with a fine which may extend up to two lakh rupees, or with both, shall be punished with imprisonment for a term which may extend to two–three years and with fine.

5.12.3 Punishment for Sending Offensive Messages through Communication Service

Any person who sends offensive messages using a computer resource or a communication device, or who passes any information or electronic mail which he knows is false, but for the purpose of causing annoyance, inconvenience, danger, obstruction, insult, injury, criminal intimidation, enmity, hatred, or ill-will using a computer resource or a communication device, shall be punished with imprisonment for a term which may extend to two–three years and with fine.

5.12.4 Punishment for Dishonestly Receiving Stolen Computer Resources or Communication Devices

A person who dishonestly receives or retains any stolen computer resource or communication device, knowing or having reason to believe the same to be stolen computer resource or communication device, shall be either imprisoned with a term which may extend to three years, or fined with an amount which may extend to one lakh rupees, or with both.

5.12.5 Punishment for Identity Theft

A person who fraudulently or dishonestly uses electronic signatures, passwords, or any other unique identification feature of any other person, shall be imprisoned for a term which may extend to three years and also be liable to pay a fine which may extend to rupees one lakh.

5.12.6 Punishment for Cheating by Personation by Using Computer Resource

If a person using any communication device or computer resource cheats by personation, he/she shall be imprisoned for a term which may extend to three years and also be liable to pay a fine which may extend to rupees one lakh.

5.12.7 Punishment for Violation of Privacy

A person who intentionally or knowingly captures, publishes, or transmits the image of a private area of any person without his or her consent, under circumstances violating the privacy of that person, shall be imprisoned for a term which may extend to three years, or fined with an amount not exceeding two lakh rupees, or with both.

Definitions

Transmit It means to electronically send a visual image with the intent that it be viewed by a person or persons.

Capture It refers to the act of filming or recording an image.

Publish It refers to the reproduction of say, an image, in printed or electronic form, and making it available to the public.

5.12.8 Publishing of Information which is Obscene in Electronic Form

A person who publishes or transmits or causes to be published in the electronic form, any material which is lascivious (feeling or revealing an overt sexual interest or desire) or appeals to prurient (having or encouraging an excessive interest in sexual matters, especially the sexual activity of others) interests, or if its effects to deprave (make immoral) and corrupt persons who are likely, having regard to all relevant circumstances, to read, see, or hear the matter contained or embodied in it, shall be punished.

On the first conviction, he/she shall be imprisoned for a term which may extend to five years and fined with an amount which may extend to one lakh rupees. For the second or subsequent conviction, the person shall be imprisoned for a term which may extend to ten years and fined with an amount which may extend to one lakh rupees.

5.12.9 Power of Controller to Give Directions

The Controller may, by order, direct a certifying authority or any employee of such authority to take measures or cease carrying on such activities as specified in the order if those are required to comply with the provisions of this Act, rules, or any regulations made thereunder. Moreover, any person who fails to comply with the order shall be guilty of an offence and shall be liable to be imprisoned for a period not exceeding three years, or fined with an amount not exceeding two lakh rupees, or to both.

5.12.10

If the con-
ment to
taken in
with for
offence.

The s
facilitie
fails to
to sever

5.12.

The ap
system
allowe

Any
term v

5.12

Whoe
fying
which
both.

5.1

If ar
tion.
tron
pers
amo

5.1

No
per

1.
2.
3.

Ar
an

5.12.10 Directions of Controller to Subscriber to Extend Facilities to Decrypt Information

If the controller is satisfied that it is necessary, he/she may, by order, direct any agency of the government to intercept any information transmitted through any computer resource. This decision shall be taken in the interest of the sovereignty or integrity of India, the security of the state, friendly relations with foreign states or public order, or for preventing incitement to the commission of any cognizable offence.

The subscriber or any person in charge of the computer resource shall, when called upon, extend all facilities and technical assistance to decrypt the information. If the subscriber or any person in charge fails to assist the agency who called him/her, he/she shall be imprisoned for a term which may extend to seven years.

5.12.11 Protected System

The appropriate government may, by notification in the official gazette, declare any computer, computer system, or computer network to be a protected system. The government may authorize persons who are allowed to access those protected systems.

Any person who either accesses or attempts to access a protected system shall be imprisoned for a term which may extend to ten years and also be liable to pay a fine.

5.12.12 Penalty for Misrepresentation

Whoever makes any misrepresentation or suppresses any material fact from the controller or the certifying authority for obtaining any licence or digital signature certificate shall be imprisoned for a term which may extend to two years, or fined with an amount which may extend to one lakh rupees, or with both.

5.12.13 Breach of Confidentiality and Privacy

If any person accesses a protected system, electronic record, book, register, correspondence, information, document, or other material without the consent of the person concerned and discloses such electronic record, book, register, correspondence, information, document, or other material to any other person, he/she shall be either imprisoned for a term which may extend to two years, or fined with an amount which may extend to one lakh rupees, or with both.

5.12.14 Penalty for Publishing Digital Signature Certificate Falsely in Certain Particulars

No person shall publish a digital signature certificate or otherwise make it available to any other person if

1. the certificate has not been issued;
2. the subscriber listed in the certificate has not accepted it; or
3. the certificate has been revoked or suspended.

Any person who does so shall be imprisoned for a term which may extend to two years, or fined with an amount which may extend to one lakh rupees, or with both.

5.12.15 Publication for Fraudulent Purpose

A person who knowingly creates, publishes, or makes available a digital signature certificate for any fraudulent or unlawful purpose shall be imprisoned for a term which may extend to two years, or fined with an amount which may extend to one lakh rupees, or with both.

5.12.16 Act to Apply for Offence or Contravention Committed outside India

The provisions of this Act shall apply also to any offence committed outside India by any person irrespective of his/her nationality.

Moreover, the provisions of this Act shall also be applicable to an offence committed outside India and by any person irrespective of his/her nationality if the Act or conduct involves a computer, computer system, or computer network located in India.

5.12.17 Confiscation

Any computer, computer system, electronic storage device (compact disks, tape drives, etc.) or any other accessory that has been contravened shall be liable to confiscation (the action of taking or seizing someone's property with authority).

5.12.18 Penalties and Confiscation Not to Interfere with Other Punishments

No penalty imposed or confiscation made under this Act shall prevent the imposition of any other punishment.

5.12.19 Power to Investigate Offence

A police officer not below the rank of Deputy Superintendent of Police shall investigate any offence under this Act.

Summary

- The Government of India enacted the Information Technology (IT) Act to deliver and facilitate lawful electronic, digital, and online transactions, and mitigate cybercrimes.
- The IT Act defines the offences in a detailed manner along with the penalties for each category of offence.
- The IT Act contains 13 chapters and 90 sections.
- The Central government may appoint a controller of certifying authorities and a number of deputy controllers and assistant controllers as it deems fit for the purposes of this Act.
- Digital signature certificate shall not be suspended for a period exceeding fifteen days unless the subscriber has been given an opportunity of being heard in the matter.
- When a digital signature certificate is issued, the public key of the subscriber is mentioned in the certificate. The subscriber has to choose a private key.
- If a person who was expected to maintain the book of accounts or records fails to do the same, he/she shall be liable to a penalty not exceeding ten thousand rupees for every day during which the failure continues.

- To adjudge whether any person has committed a contravention of any of the provisions of this Act, the Central government shall appoint any officer not below the rank of a director to the Government of India.
- The Central government shall, establish one or more appellate tribunals to be known as the Cyber Regulations Appellate Tribunal.
- The Presiding Officer of Cyber-Appellate Tribunal shall hold office for a term of five years from the date on which he/she enters upon his/her office or until he/she attains the age of sixty five years, whichever is earlier.
- The person who knowingly or intentionally conceals, destroys, or alters any computer source code used for a computer, computer program, computer system, or computer network, when the source code is required to be kept or maintained by law for the time being in force, shall be imprisoned for three years, or fined with an amount which may extend up to two lakh rupees, or be punished with both.
- A person who dishonestly receives or retains any stolen computer resource or communication device knowing or having reason to believe the same to be a stolen computer resource or communication device, shall be either imprisoned with a term which may extend to three years, or fined with an amount which may extend to one lakh rupees, or with both.
- Any person who either accesses or attempts to access a protected system shall be imprisoned for a term which may extend to ten years and also be liable to pay a fine.
- A person who knowingly creates, publishes, or makes available a digital signature certificate for any fraudulent or unlawful purpose shall be imprisoned for a term which may extend to two years, or fined with an amount which may extend to one lakh rupees, or with both.
- Some common examples of cybercrimes include identity theft, financial fraud, website defacements, cyber bullying, hacking of databases, theft of intellectual property, cyber harassment (harassing a person based on sexual, racial, religious, or other factors), software piracy, transmission of harmful programs, and unauthorized access of computerized information.

Glossary

Auditor Any internationally accredited computer security professional or agency appointed by the certifying authority and recognized by the controller for doing a technical audit of the work done by the certifying authority

Capture To film or record an image

Computer database A representation of information, knowledge, facts, concepts, or instructions in text, image, audio, video that are being prepared, or have been prepared in a formalized manner, or have been produced by a computer, computer system, or computer network and are intended for use in a computer, computer system, or computer network

Computer virus Any computer instruction, information, data, or program that destroys, damages, degrades, adversely affects the performance of a

computer resource or attaches to another computer resource and operates when a program, date, or instruction is executed or some other event takes place in that computer resource

Cyber café Any facility from where the Internet access is offered by any person in the ordinary course of business to the members of the public

Damage Destroy, alter, delete, add, modify, or re-range any computer resource by any means

Digital signature A technique that is used to validate the legitimacy of a digital message or a document. A valid digital signature ensures the recipient that the message was generated by a known sender

Information asset Refers to all information resources utilized in the organization for conducting its business. It includes all the information, applications

(i.e., software developed or purchased), and technology (hardware, system software, and networks)

Licence A licence (or permission) granted to certifying authorities for the issue of digital signature certificates under the defined rules

Licensed certifying authority The certifying authority who has the licence to issue digital signature certificates

Person An individual, or a company, or an association, or a body of individuals, or Central government or a state government, or any of the ministries or departments, agencies, or authorities of such governments

Publishing Reproducing say an image, in the printed or electronic form and making it available for public

Transmit Electronically send a visual image with the intent that it be viewed by a person or persons

Multiple-choice Questions

- Who issues the digital certificate?
 - Controller
 - Auditor
 - Certifying authority
 - Applicant
- How many Schedules are there in the IT Act 2000?
 - 3
 - 4
 - 6
 - 2
- What is/are the component(s) of the IT Act 2000?
 - Legal recognition to digital signatures
 - Regulation of certification authorities
 - Digital certificates
 - All of these
- Under whom does the controller of certifying authorities work?
 - Prime Minister's Office
 - Reserve Bank of India
 - Ministry of Communication & IT
 - Autonomous body
- What is the penalty for publishing images of a person's private parts without consent, as per the IT Act 2000?
 - Five years imprisonment or 5 lakh rupees penalty, or both
 - Life imprisonment
 - Three years imprisonment or 2 lakh rupees penalty, or both
 - None of these
- What is the time limit for filing an appeal against the order of the Cyber Appellate Tribunal?
 - 30 days
 - 90 days
 - 60 days
 - 45 days
- What is the punishment for hacking computers?
 - Three years imprisonment or 10 lakh rupees penalty, or both
 - Life imprisonment
 - Three years imprisonment or 5 lakh rupees penalty, or both
 - Three years imprisonment and 5 lakh rupees penalty
- What is the term of the office of the Presiding Officer of the Cyber Appellate Tribunal?
 - 3 years
 - 4 years
 - 6 years
 - 5 years
- Which of the following is the court for appealing against the orders issued by the Cyber Appellate Tribunal?
 - Family court
 - District court
 - High court
 - Supreme Court
- Which of the following is an example of intellectual property?
 - Trademarks
 - Copyright
 - Patent
 - All of these
- Transactions carried out using electronic data interchange, and other means of communication are known as _____.
 - e-communication
 - e-commerce

- (b) e-governance (d) e-filing
12. The authentication to be affected by using asymmetric encryption and hash function is known as _____.
- (a) public key (c) digital signature
(b) private key (d) e-governance
13. An electronic record is verified using _____.
- (a) public key (c) digital signature
(b) private key (d) e-governance
14. A digital signature is created using _____.
- (a) Cryptography (c) Program
(b) HTML (d) Coding
15. A public key is used to _____.
- (a) sign an electronic document
(b) verify a signature
(c) sign digitally
(d) make payments
16. The Central government appoints a _____ who supervises the work of the certifying authority.
- (a) controller (c) subscriber
(b) commissioner (d) superintendent
17. The digital certificate is issued in the name of the _____.
- (a) controller (c) subscriber
(b) commissioner (d) superintendent
18. A digital signature is issued by the _____.
- (a) Central government
(b) state government
(c) commissioner of certifying authority
(d) controller of certifying authority
19. A digital certificate can be revoked by the _____.
- (a) Central government
(b) state government
(c) certifying authority
(d) subscriber
20. Which of the following is not a part of e-governance?
- (a) Use of e-records
(b) Use of e-gazette
(c) Recognition of digital signature
(d) None of these
21. The controller works under the directions of the _____.
- (a) state government
(b) Central government
(c) certifying authority
(d) Supreme Court
22. Who lays out the format and content of a digital signature certificate and the key?
- (a) State government
(b) Central government
(c) Controller of certifying authority
(d) Supreme Court
23. The fees to apply for a digital signature should not exceed ₹ _____.
- (a) 10,000 (c) 30,000
(b) 25,000 (d) 50,000
24. The private key is chosen by the _____.
- (a) subscriber
(b) state government
(c) Central government
(d) controller of certifying authority
25. A digital signature certificate shall not be suspended for a period exceeding _____.
- (a) 1 week (c) 1 month
(b) 15 days (d) 1 year
26. If the certifying authority finds that a material fact represented in the digital signature certificate is false or has been concealed, it will _____ the certificate.
- (a) suspend (c) cancel
(b) revoke (d) issue
27. If the certifying authority's security system has been compromised in a way that can affect the reliability of the digital signature certificate, it _____ all certificates.
- (a) suspends (c) cancels
(b) revokes (d) issues
28. In a digital certificate, _____ key of the subscriber is mentioned.
- (a) public (c) symmetric
(b) private (d) session
29. The penalty for damaging a computer or network can go up to ₹ _____.

- (a) 1 lakh (c) 1 crore
(b) 10 lakhs (d) 10 crore
30. The penalty for failure to furnish information cannot exceed ₹ _____.
(a) 50,000 (c) 1.5 lakh
(b) 1 lakh (d) 2 lakh
31. Inquiries in case of adjudging contraventions are made by the _____.
(a) Controller of certifying authority
(b) Director
(c) Police Commissioner
(d) Adjudging Officer
32. The officers in the Cyber Appellate Tribunal work under the supervision of _____.
(a) Controller of certifying authority
(b) Presiding Officer
(c) Police Superintendent
(d) High court judge
33. Any person aggrieved by an order made by the controller or an adjudicating officer may appeal to the _____.
(a) High court
(b) Certifying authority
(c) Cyber Appellate Tribunal
(d) Central government
34. Any person can file an appeal to the Cyber Appellate Tribunal within _____ days of order passed.
(a) 30 (c) 60
(b) 45 (d) 90
35. The appeal filed before the Cyber Appellate Tribunal shall be dealt within _____ from the date of receipt of the appeal.
(a) 6 months (c) 2 years
(b) 1 year (d) 3 years
36. A person can file an appeal in _____ days against the orders of the Cyber Appellate Tribunal.
(a) 30 (c) 60
(b) 45 (d) 90
37. The penalty for tampering with computer documents is _____.
(a) imprisonment for three years, or fine with an amount which may extend up to two lakh rupees, or both
(b) imprisonment up to three years, or with fine which may extend up to two lakh rupees, or both
(c) imprisonment for a term which may extend to two–three years and with fine
(d) imprisonment with a term which may extend to three years or fine with an amount which may extend to rupees one lakh, or both
38. The penalty for hacking a computer system is _____.
(a) imprisonment for three years, or fine with an amount which may extend up to two lakh rupees, or both
(b) imprisonment up to three years, or with fine which may extend up to two lakh rupees, or both
(c) imprisonment for a term which may extend to two–three years and with fine
(d) imprisonment with a term which may extend to three years, or fine with an amount which may extend to rupees one lakh, or both
39. The penalty for sending offensive messages is _____.
(a) imprisonment for three years, or fine with an amount which may extend up to two lakh rupees, or both
(b) imprisonment up to three years, or fine which may extend up to two lakh rupees, or both
(c) imprisonment for a term which may extend to two–three years and with fine
(d) imprisonment with a term which may extend to three years, or fine with an amount which may extend to rupees one lakh, or with both
40. The penalty for dishonestly receiving a stolen computer is _____.
(a) imprisonment for three years, or fine with an amount which may extend up to two lakh rupees, or both

- (b) imprisonment up to three years, or fine which may extend up to two lakh rupees, or both
- (c) imprisonment for a term which may extend to two-three years and with fine
- (d) imprisonment with a term which may extend to three years, or fine with an amount which may extend to rupees one lakh, or both
41. The penalty for identity theft is _____
- (a) imprisonment for three years, or fine with an amount which may extend up to two lakh rupees, or both
- (b) imprisonment up to three years or fine which may extend up to one lakh rupees
- (c) imprisonment for a term which may extend to two-three years and with fine
- (d) imprisonment with a term which may extend to three years, or fine with an amount which may extend to rupees one lakh, or both
42. The penalty for cheating by personification by using a computer resource is _____
- (a) imprisonment for three years, or fine with an amount which may extend up to two lakh rupees, or both
- (b) imprisonment up to three years or fine which may extend up to one lakh rupees
- (c) imprisonment for a term which may extend to two-three years and with fine
- (d) imprisonment with a term which may extend to three years, or fine with an amount which may extend to rupees one lakh, or both
43. The penalty for violation of privacy is _____
- (a) imprisonment for three years, or fine with an amount which may extend up to two lakh rupees, or both
- (b) imprisonment up to three years or fine which may extend up to one lakh rupees
- (c) imprisonment for a term which may extend to two-three years and with fine
- (d) imprisonment with a term which may extend to three years, or fine with an amount which may extend to rupees one lakh, or both
44. The penalty for the first contravention in case of publishing obscene information in electronic form is _____
- (a) imprisonment for a period which may extend to five years and fine with an amount which may extend up to one lakh rupees
- (b) imprisonment for a period which may extend to 10 years and fine with an amount which may extend up to one lakh rupees
- (c) imprisonment up to three years or with fine which may extend up to one lakh rupees
- (d) imprisonment for a term which may extend to two-three years and with fine
45. The penalty for the second contravention in case of publishing obscene information in electronic form is _____
- (a) imprisonment for a period which may extend to five years and fine with an amount which may extend up to one lakh rupees
- (b) imprisonment for a period which may extend to 10 years and fine with an amount which may extend up to one lakh rupees
- (c) imprisonment up to three years or fine which may extend up to one lakh rupees
- (d) imprisonment for a term which may extend to two-three years and with fine
46. Any person who either accesses or attempts to access a protected system shall be penalized with _____
- (a) imprisonment for a term which may extend to 10 years and also be liable to pay a fine
- (b) imprisonment for a period which may extend to 10 years and fine with an amount which may extend up to one lakh rupees
- (c) imprisonment up to three years or with fine which may extend up to one lakh rupees

- (d) imprisonment for a term which may extend to two–three years and with fine
47. Whoever makes any misrepresentation or suppresses any material fact from the Controller or the certifying authority for obtaining any licence or digital signature certificate shall be penalized with _____.
- (a) imprisonment for a term which may extend to 10 years and also be liable to pay a fine
- (b) imprisonment for a term which may extend to two years, or fine with an amount which may extend to one lakh rupees, or both
- (c) imprisonment up to three years or fine which may extend up to one lakh rupees
- (d) imprisonment for a term which may extend to two–three years and with fine
48. In case of breach of confidentiality and privacy, the penalty is _____.
- (a) imprisonment for a term which may extend to 10 years and also be liable to pay a fine

- (b) imprisonment for a term which may extend to two years, or fine with an amount which may extend to one lakh rupees, or both
- (c) imprisonment up to three years or fine which may extend up to one lakh rupees
- (d) imprisonment for a term which may extend to two–three years and with fine
49. The penalty for using a digital certificate that has been revoked is _____.
- (a) imprisonment for a term which may extend to 10 years and also be liable to pay a fine
- (b) imprisonment for a term which may extend to two years, or fine with an amount which may extend to one lakh rupees, or with both
- (c) imprisonment up to three years or with fine which may extend up to one lakh rupees
- (d) imprisonment for a term which may extend to two–three years and with fine

Answers to Multiple-choice Questions

- | | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1. (c) | 2. (b) | 3. (d) | 4. (c) | 5. (c) | 6. (c) | 7. (c) | 8. (d) | 9. (c) | 10. (d) |
| 11. (c) | 12. (c) | 13. (c) | 14. (a) | 15. (b) | 16. (a) | 17. (c) | 18. (d) | 19. (c) | 20. (d) |
| 21. (b) | 22. (c) | 23. (b) | 24. (a) | 25. (b) | 26. (b) | 27. (b) | 28. (a) | 29. (c) | 30. (c) |
| 31. (d) | 32. (b) | 33. (c) | 34. (b) | 35. (a) | 36. (c) | 37. (a) | 38. (b) | 39. (c) | 40. (d) |
| 41. (b) | 42. (b) | 43. (d) | 44. (a) | 45. (b) | 46. (a) | 47. (b) | 48. (b) | 49. (b) | |

Elementary Word

Working with check, autocorrect, indent, page numbering, and formatting with data facilities.

6.1 INTRODUCTION

Word 2013 is a Microsoft Office application.

- Create and share documents
- Create and manage templates
- Generate reports
- Design and format documents
- Make and manage lists
- Create and manage tables

6.2 OPENING A DOCUMENT

To begin working with Word, you need to open a document, or file.

A file is a document.

Step 1: Click the File tab.

Step 2: Click Open.

Step 3: Select the document you want to open.

Step 4: Click Open.

A new document will be opened.

the Title bar will appear.



Elements of Word Processing

Syllabus Mapping

Unit

Working with word document—editing text, find and replace text, formatting, spell check, autocorrect, auto text; bullets and numbering, tabs, paragraph formatting, indent, page formatting, header and footer, macros, drop cap; tables—inserting, filling, and formatting a table, inserting pictures and video; mail merge—including linking with database, printing documents; creating business documents using the above facilities.

Module II
Unit 6

6.1 INTRODUCTION

Word 2013 is a word processing application software developed by Microsoft. It comes as a part of Microsoft Office 2013 suite. You can do a lot on MS Word:

- Create professional-looking documents with different themes, visual designs, formatting tools, and sharing features.
- Create and edit personal and business documents including letters, invoices, e-mails, and books.
- Generate business reports having various graphics including pictures, charts, and diagrams.
- Design flyers, cover pages, and letter heads for personal and business purpose.
- Make documents such as résumés or invitation cards.
- Create a wide range of documents varying from a simple office memo to legal copies and reference documents.

6.2 OPENING, PRINTING, AND CLOSING DOCUMENTS

To begin working with MS Word let us now try to perform basic operations such as creating a new document, opening an existing document, saving a document, etc.

A file created using MS Word is called a document. Let us create a new document.

Step 1: Click on the File tab.

Step 2: Click on New.

Step 3: Click on Blank Document from the list of options as shown in Fig. 6.3.

Step 4: Click on Create.

A new document is created and Word names it as Document 1. You can see that this name appears in the Title bar. Now you can type your text in the document area.



To create a new document you can also press Ctrl + N keys on the keyboard.

6.2.1 Opening Documents

Step 1: Click on the File tab.

Step 2: Click on Open. The Open dialog box appears.

Step 3: Select File and click on Open.

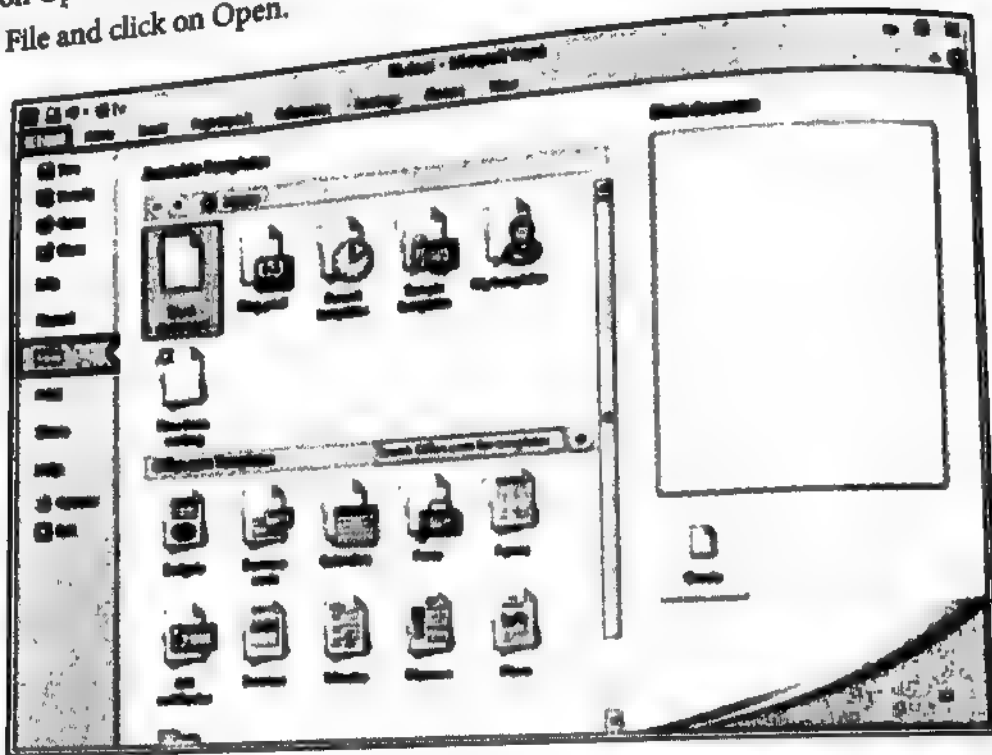
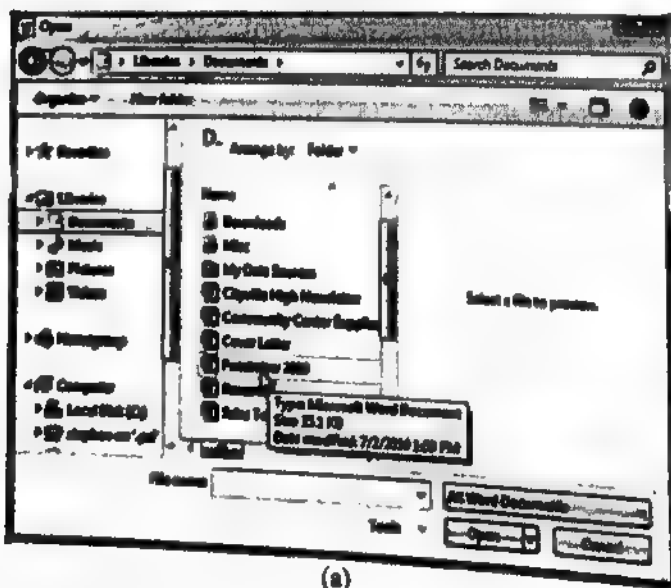
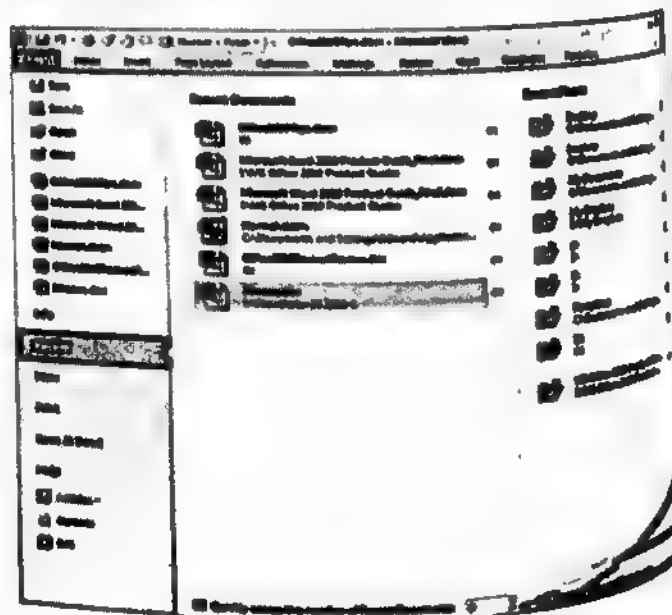


Figure 6.1 Creating a new document

There is another way of opening a file. If you want to open a file that had been used recently, then you can click on *Recent* option in the File menu. A list of recently opened Word files is displayed. Click on the name of the file that you want to open as shown in Fig. 6.2.



(a)



(b)

Figure 6.2 Opening Word document

6.2.2 Saving Documents

After creating a new document, you need to save it. The first time you save a document, you are prompted to enter a name for the file and an extension.

Step 1: Click on the File tab.

Step 2: Click on Save.

Step 3: Enter a name for the file.

Step 4: Click on Save.

Once you have saved a document, you can open it whenever you need it. You must save the document before you change it.

To save a document, you need to enter a name for the file.



6.2.3 Printing Documents

The first time you print a document, you are prompted to enter a name for the file and an extension.

Step 1: Click on the File tab.

Step 2: Click on Print.

Step 3: Enter a name for the file.

Step 4: Click on Print.

Once you have printed a document, you can open it whenever you need it. You must save the document before you change it.

To save a document, you need to enter a name for the file.

Step 1: Click on the File tab.

Step 2: Click on Save.

Step 3: Enter a name for the file.

Step 4: Click on Save.

Once you have saved a document, you can open it whenever you need it. You must save the document before you change it.

To save a document, you need to enter a name for the file.

6.2.2 Save and Save As

After creating a new document, you must immediately save it. This will save your document permanently on the hard disk. You must save your document with a valid name. After saving the document, the name of the document will appear in the Title bar. In MS Word, a document is automatically saved with an extension .docx. To save a document, follow the following steps:

Step 1: Click on File tab.

Step 2: Click on Save As option.

Step 3: Select a folder where you would like to save the document and enter the name you want to give to save your document, as shown in Fig. 6.3.

Step 4: Click on Save button.

Once the document is saved, you can always open it for editing or inserting new text. But whenever you make any changes in it, you must save it again. Even while working on the file, you must save your work every few minutes so that changes are made permanently in the file.

To save your changes under the same file name, you can either press Ctrl + S keys or click on the Save icon in the File tab. The Save icon is also present on the Quick Access Toolbar.

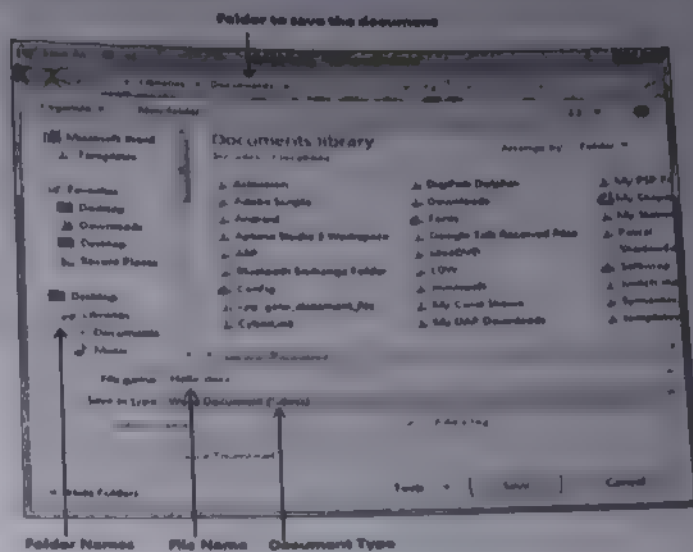


Figure 6.3 Saving a document



While typing, press Enter key only when you want to start a new paragraph.

6.2.3 Page Setup

The Page Setup dialog box has the following three tabs:

Margins This tab is used for setting margins. You can apply margins either on the whole document or a part of it. You can set the following margins:

Top It specifies the distance from the top of the page to the first line of the text.

Bottom It is used to specify the distance from the bottom of the page to the last line of text.

Left It specifies the distance from the left edge of the page to the start of the text.

Right It adjusts the distance from the right edge of the page to the end of the text.

Gutter It specifies an additional space of margin for pages that are to be bound.

Paper This tab is used to select the paper size for all the pages in your document. It also allows you to specify the printer tray to use for the first and other pages of the Word document. For example, this is an important feature when you want to use a heavier paper or a coloured paper for the title page and white paper for the remainder of your document.

Layout In this tab, formatting features that do not fit in the other tabs are specified. For example, line numbers, headers, footers, page alignment, borders, etc.

To open the Page Setup dialog box follow the steps given below.

Step 1: Click on the Page Layout tab.

Step 2: In the Page Setup group, click on the small arrow in the bottom-right corner.

Step 3: The Page Setup dialog box (Fig. 6.4) will appear. Make the desired changes and click on OK button to confirm the changes.

Page margins are the blank space around the edges of the page. Text and graphics are inserted in the printable area between the margins. Header, footers, and page numbers can be placed in the margins.

From the Margin tab, you can also choose the orientation of the page (*landscape*—vertical or *portrait*—horizontal) and the part of the document on which the margin has to be applied—whether whole document (*Whole Document*), from the insertion pointer's position onward (*This Point Forward*), only on the selected text (*Selected Text*) or only to the current section (*This Section*). The *Multiple Pages* drop-down list tells Word how to print the document on a page.

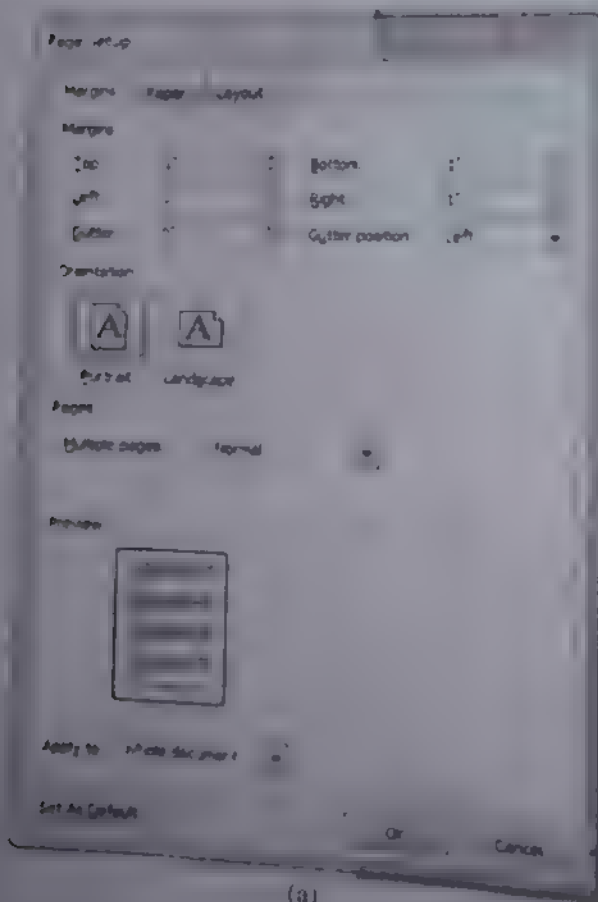


To open the Page Setup dialog box, you can also press (Alt + P) followed by S, followed by P.

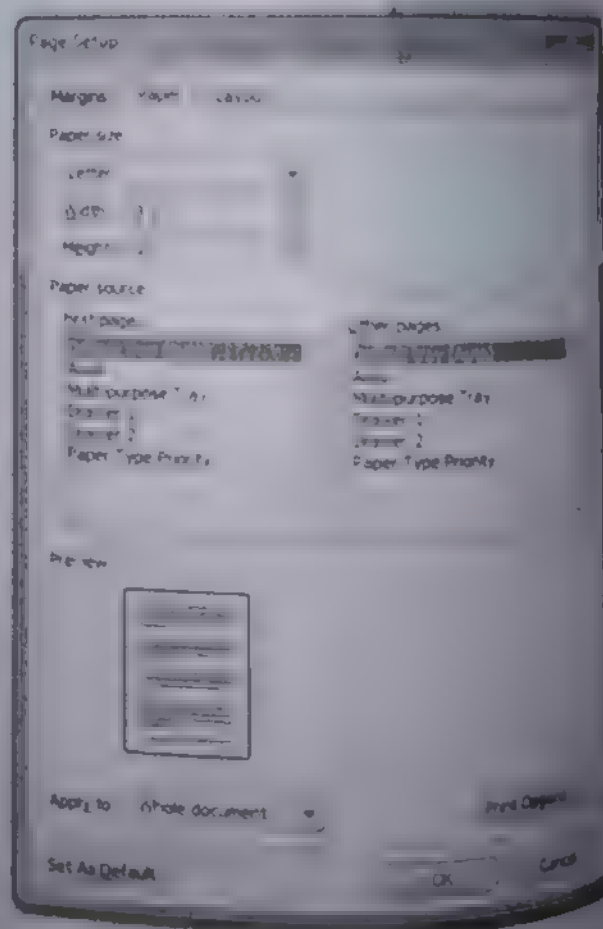
6.2.4 Print Preview, Printing and Closing Documents

To print and preview files, follow the steps given below:

Step 1: Click on File tab.



(a)

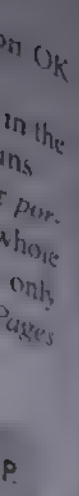


(b)

Step 2: Click on the Preview button.

In the Preview window, you can see the preview of the document. You can look where the margins are and make changes between the margins.

Step 3: Click on the Print button.



(c)

Figure 6.4 Page Setup dialog box

Step 2: Click on Print as shown in Fig. 6.5.

In the window that pops up, the properties for your default printer is shown in the first section, and the preview of the document is seen in the second section. Print Preview is used to see how a page would look when printed on a page. The Print Preview section has page navigation buttons to help users move between pages of a document. It also has a zoom slider to zoom in and zoom out. Whenever you make a change to a print-related setting, the preview is automatically updated.

Step 3: If the properties for your printer and document are fine with you, click on Print to print the document.

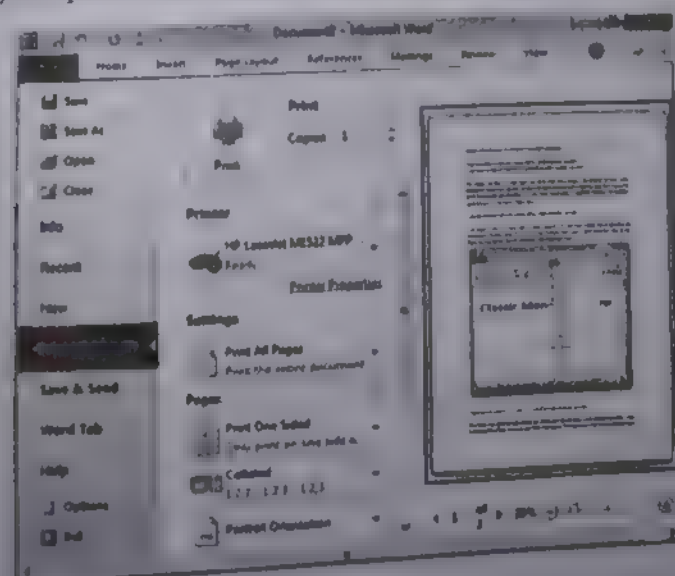


Figure 6.5 Print window

Printing a Part of a Document

Before printing a document, you can also make the following choices:

Print all or part of your document. Under Settings click on Print All Pages to choose what to print. For example,

- Select *Print All Pages* to print the entire document.
- Select *Print Selection* to print only the selected content.
- Select *Print Current Page* to print only the current page.
- Select *Print Custom Range* to print only some pages in the document. In the Pages box, you can enter the page numbers and or page ranges separated by commas. For example, type 2, 5, 6-9
- Select *Only Print Odd Pages* to print odd pages in the document and *Only Print Even Pages* to print all even pages in the document.

Closing a Document

To close a document, click on the Close button. You can also press the Alt + F4 keys. However, before closing the document make sure that you have saved your work.

6.3 TEXT CREATION AND MANIPULATION

In this section, we will learn about entering and formatting text in MS Word. Entering text means typing text and formatting means to apply some styles to the text to make it more presentable and attractive.

6.3.1 Document Creation

In the last section, we studied that to create a document either press Ctrl + N keys or click on File tab followed by New option. Once the document is created type some text. The text would be inserted at the position of the cursor.

Inserting text means adding a word or text into the existing content without replacing any existing content in the file. To insert text, follow the steps given below:

Step 1: Click on the location where you want to insert text. You can even use arrow keys on the keyboard to move to the position where you would like to insert the text.

Step 2: Type the text you want to insert. The new text would be inserted to the left of the insertion point.



If you want to add more than one space between two words, use the Tab key.

6.3.2 Text Editing

Editing a file means making some changes in the file. In Word, editing a file is as simple as entering text in a blank document. While editing a document we may insert, delete, or replace some existing content. Inserting text has been discussed in the previous section. So in this section, we will see how to replace and delete text from a file.

There are multiple ways to delete. You can simply highlight the text to be deleted using your mouse and press the Delete key.

To delete only position and then to delete the character before the character being typed. If you want to delete the new one.

MS Word also has an Undo button to reverse the changes.

6.3.3 Text

Selecting a text can be performed in several ways. The simplest is to click on the text. Similarly, use the mouse to select text (holding down the mouse button).

To select text (holding down the mouse button), hold down the mouse button and drag the mouse. Press F8 and then the arrow keys.

- To select text (holding down the mouse button), hold down the mouse button and drag the mouse.
- To select text (holding down the mouse button), hold down the mouse button and drag the mouse.
- To select text (holding down the mouse button), hold down the mouse button and drag the mouse.

You can also use the figure is an arrow. Now

To delete only some character(s), place the cursor at the appropriate position and then press Backspace key or Delete key. Press the Delete key to delete the character after the cursor. Press the Backspace key to delete the character before the cursor.

If you want to replace a text with another text, then highlight the text and type the new text. The existing text will be replaced (overwritten) by the new one.

MS Word also has Undo and Redo buttons as shown in Fig. 6.6. The Undo button is used to reverse the last action(s) performed. Redo is used to reverse the effect of Undo button.

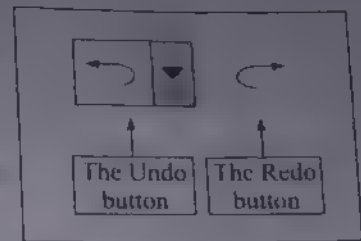


Figure 6.6 Undo and Redo buttons

6.3.3 Text Selection

Selecting a text is one of the most important things that we do in a Word document. Many operations can be performed on the selected text. Like, you can delete it, copy it, cut it, format it, align it, indent it, and overwrite it, etc.

The simplest way to select text is to click on it and drag the mouse over the text you want to select. Similarly, use the following techniques to select text.

To select text between two points, click on the start of the text (from where you want to start selecting), hold down the Shift key, and click at the end of the text (where you want to stop selecting), OR

Holding down the Shift key, use any of the arrow keys to select the text, OR
Press F8 and then use any of the arrow keys to select the text.

- To select a single word, double click on the text.
- To select a paragraph, click three times on the paragraph.
- To select the whole document, press Ctrl + A keys on the keyboard.

You can even use the Selection bar to select a text as shown in Fig. 6.7. The black shaded area in the figure is called *selection bar*. Bring your cursor in this area. It would turn into a rightward-pointing arrow. Now you can select using the following techniques.

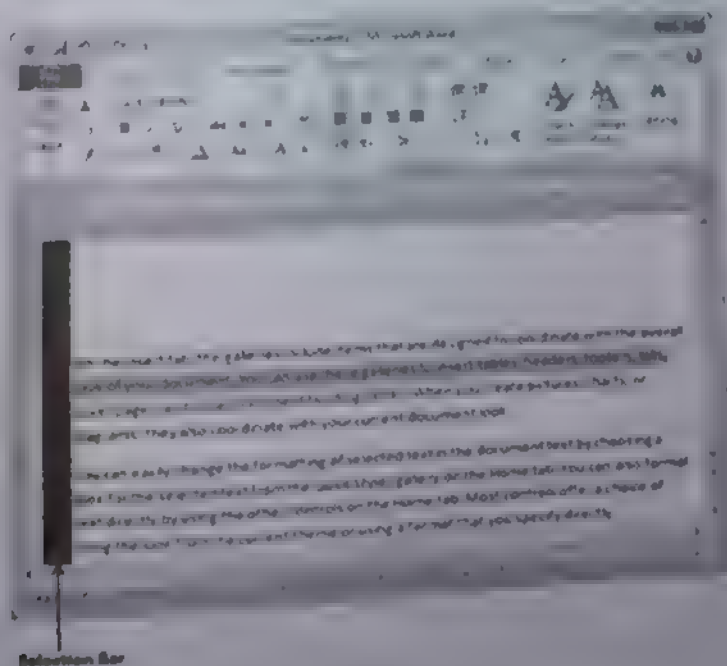


Figure 6.7 Selecting text

- To select a line, click in front of the line you want to select.
- To select the whole document, click three times.

6.3.4 Cut, Copy, and Paste

The Cut Copy command must always be followed by the Paste command. While Copy Paste is used to duplicate some content at another location, the Cut Paste operation is used to delete the content from one location and insert it into another location.

Word has a small area to hold data temporarily. This area is called clipboard. When you perform Copy or Cut operations, Word stores the data temporarily in the clipboard. When you do the Paste operation, it copies the data from the clipboard and pastes it at the desired location.

Let us perform the Copy and Paste operations to copy the content from its original location and duplicate it at the desired location. Note that in copy, the contents are not deleted from their original location. Follow the steps given below to perform Copy and Paste operations.

Step 1: Select the text.

Step 2: Copy the selected text on the clipboard. This can be done in any one of the following ways.

- By right clicking on the selected text and selecting Copy option from the pop-up menu
- Clicking on Copy button available in the ribbon
- Pressing Ctrl + C keys

Step 3: Click at the place where you want to paste the selected text. This would move the cursor to the desired location.

Step 4: Perform Paste command by any of the following ways.

- Clicking on the Paste button in the Ribbon as shown in Fig. 6.8.

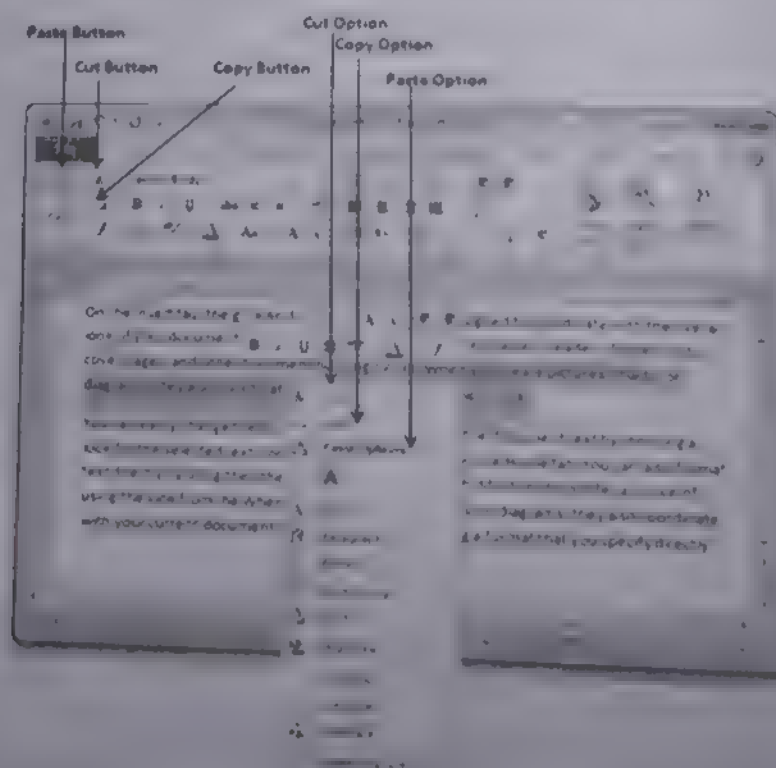


Figure 6.8 Cut, Copy, Paste commands

- Pressing Ctrl + V keys on the keyboard
 - By right clicking at the desired location and selecting paste option from the pop up menu
- To create multiple duplicate copies, you can paste as many times as you want.

Cut and Paste Operation

Cut operation will delete the content from its original place and paste will insert it at the desired location. To cut and paste in Word, follow the steps given below:

- Step 1:** Select a portion of the text.
- Step 2:** Cut the selected text. The text will be copied on the clipboard. This can be done in any one of the following ways.
- By right clicking on the selected text and selecting Cut option from the pop-up menu.
 - Clicking on Cut button available in the Ribbon
 - Pressing Ctrl + X keys.
- Step 3:** Click at the place where you want to paste the selected text. This would move the cursor to the desired location.
- Step 4:** Perform Paste command by any of the following ways.
- Clicking on the Paste button in the Ribbon.
 - Pressing Ctrl + V keys on the keyboard.
 - By right clicking at the desired location and selecting Paste option from the pop-up menu.



Cut, Copy, and Paste operations can be performed within the same document as well as across multiple documents.

6.3.5 Font and Size Selection

Word allows you to use different fonts with different sizes. This helps you to change the appearance of the text and make it more attractive. Generally, we use different fonts and sizes for headings, sub-headings, text, important terms, etc.

Follow the steps given below to change the font and its size:

- Step 1:** Select the text whose font and size has to be changed.
- Step 2:** Click on the Home tab.
- Step 3:** Click on the triangle next to Font Type to list down all the fonts available. Select a font.
- Step 4:** Click on the triangle next to Font Size to display a font size list. You can also type the size. This is shown in Fig. 6.9.

Let us also learn how to grow and shrink the size of the text. Fig. 6.10 shows these buttons. Select the text and click on Shrink button to

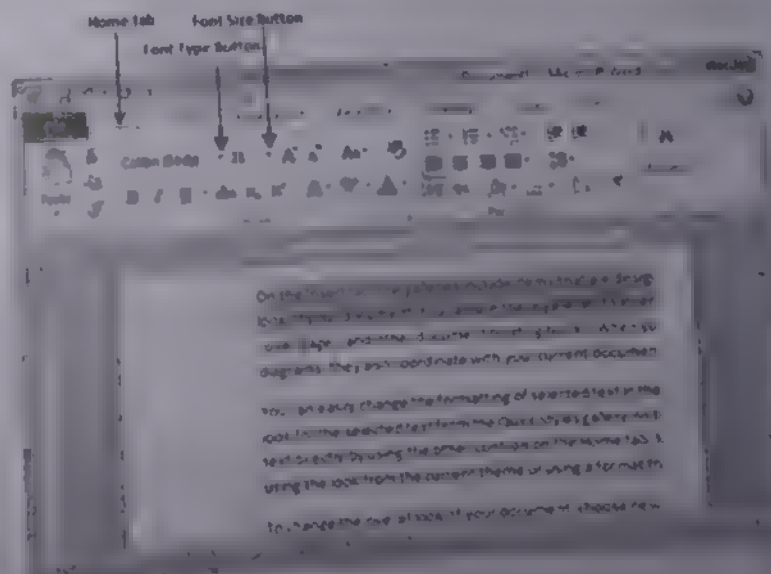


Figure 6.9 Setting the font

reduce the size and click on Grow button to enlarge the font size of the text. You can click these buttons multiple times. Each time the button is clicked, the font is enlarged or reduced by 1 point.



Figure 6.10 Increasing and decreasing font size

6.3.6 Alignment of Text

Alignment of text means how the text is displayed horizontally on the page. There are four options for text alignment – left, center, right, and justified.

Left A text is said to be left aligned if it is aligned with left margin. To left align a text, select the text and click on Left Align button, as shown in Fig. 6.11.

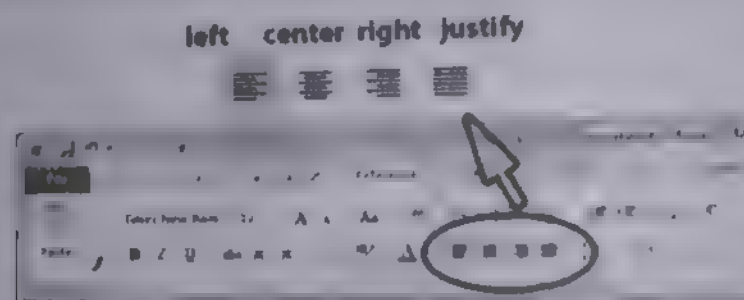


Figure 6.11 Text alignment

Right A text is said to be right aligned if it is aligned with the right margin. To right align a text, select the text and click on the Right Align button.

Justify A text is said to be justified if it is aligned with both left and right margins. To justify align a text, select the text and click on Justify Align button.

Center A text is said to be center aligned if it is center aligned between the left and right margins. To center align a text, select the text and click on Center Align button.

6.3.7 Changing Font Colours

By default, any text we type in Word comes in black colour, but you can change the font colour to any colour of your choice. Follow the steps given below to change the colour of your font:

Step 1: Select the text

Step 2: Click on the Home tab.

Step 3: Click on the triangle next to Font Colour to display a list of colours. Try to move your mouse pointer over different colours to see your selected text in different colours.

Step 4: Select a colour. If you do not find a colour of your choice, you can use More Colours option to display colour pallet box which has a wide range of colours as shown in Fig. 6.12.

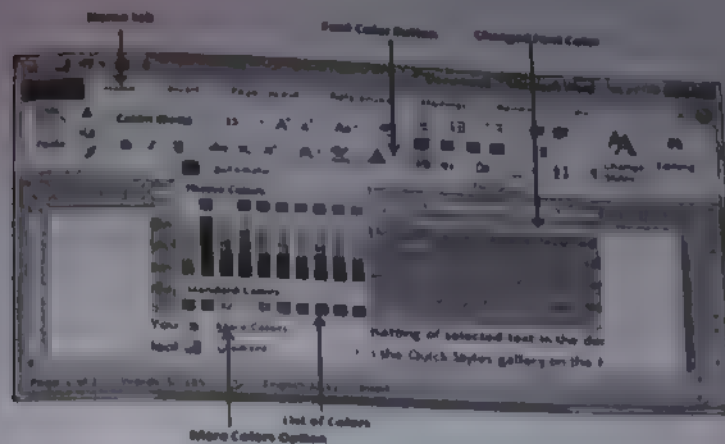


Figure 6.12 Font colour

6.3.8 Highlighting Text with Colours

To highlight a selected text using any colour, follow the following steps:

Step 1: Select the text.

Step 2: Click on the triangle next to the 'Text Highlight Colour' button to display a list of colours. When you move the mouse pointer over different colours, Text Highlight Colour will change automatically as shown in Fig. 6.13.

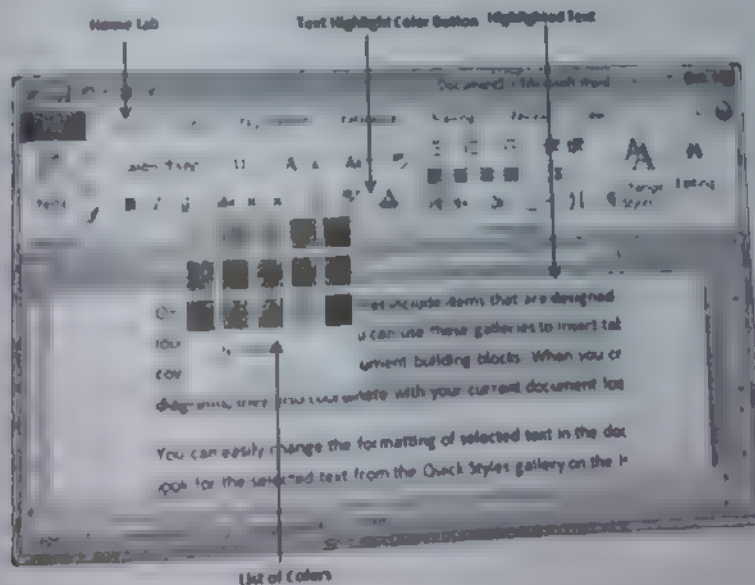


Figure 6.13 Highlighting text

Step 3: Select a colour by clicking on it.

6.3.9 Applying Text Effects

You can apply various effects on text to beautify the document. This is especially useful to design cover page or headings of the document. Text effects can be applied by following the steps given below:

Step 1: Select the text.

Step 2: Click the Text Effect button to display a list of effects including shadow, outline, glow, reflect, etc. When you move the mouse pointer over different effects, text effect will change automatically as shown in Fig. 6.14.

Step 3: Click on any of the text effects.

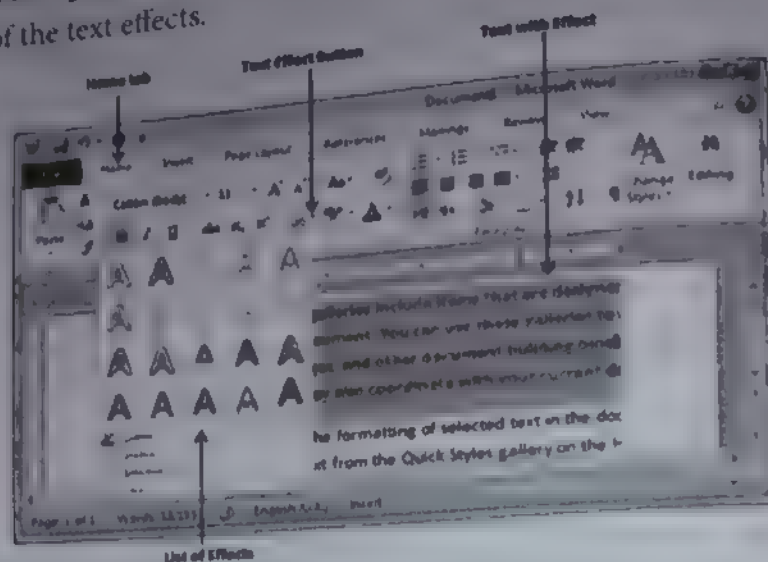


Figure 6.14 Applying text effects

6.3.10 Inserting WordArt

WordArt is decorative text that you can add to a document. Like normal text, you can also edit WordArt such as changing the font size and the text colour, by using the drawing tools options available automatically after you insert or select the WordArt in a document. WordArt is usually used to add special text effects to your document. For example, you can stretch a title, skew text, make text fit a preset shape, or apply a gradient fill. Once inserted, you can also move or position it in your document to add decoration or emphasis.

Follow the steps given below to insert WordArt in your document.

Step 1: In the document, click where you want to insert WordArt.

Step 2: Click on the Insert tab.

Step 3: From the Text group, click WordArt.

Step 4: Click on any WordArt style, and type the text as shown in Fig. 6.15.

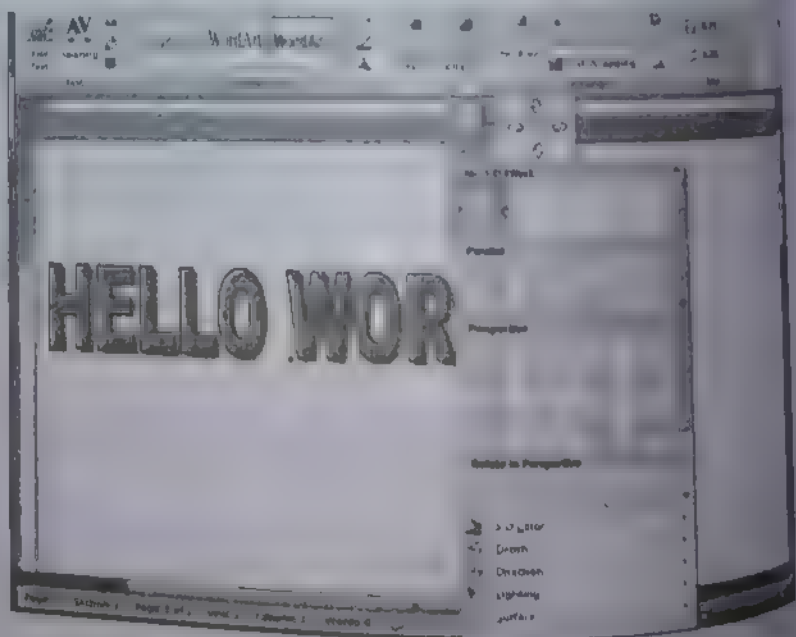


Figure 6.15 WordArt

6.4 FORMATTING THE TEXT

In this section, we will learn about some more ways in which we can format text to make it more presentable.

6.4.1 Paragraph Indenting

In the previous section, we had used margin settings to determine the blank space on each side of a paragraph. You can also indent paragraphs in a document. This indentation can be done on the left margin, the right margin, or both margins.

Left Indentation Left indentation means to move the left edge of the paragraph inward towards the center of the paragraph. To left indent a paragraph, follow the steps given below.

Step 1: Click anywhere on the paragraph.

Step 2: Click on Increase Indent button available on Home tab (refer Fig. 6.16) or press Ctrl + M keys. To have a deeper indentation you can click multiple times.

To remove left indentation, click on Decrease Indent button available on Home tab or press Ctrl + Shift + M keys. To remove deeper indentation, multiple times.

Right Indentation Right indentation means to move the right edge of the paragraph inward towards the center of the paragraph. To right indent a paragraph, follow the steps given below.

Step 1: Click anywhere on the paragraph.

Step 2: Click on Page Layout tab.

Step 3: Use the Increase Right Indent spinner or Left Indent spinners to increase the indentation of the paragraph.

Click on the Decrease Right Indent spinner as shown in Fig. 6.17) to remove indentation.

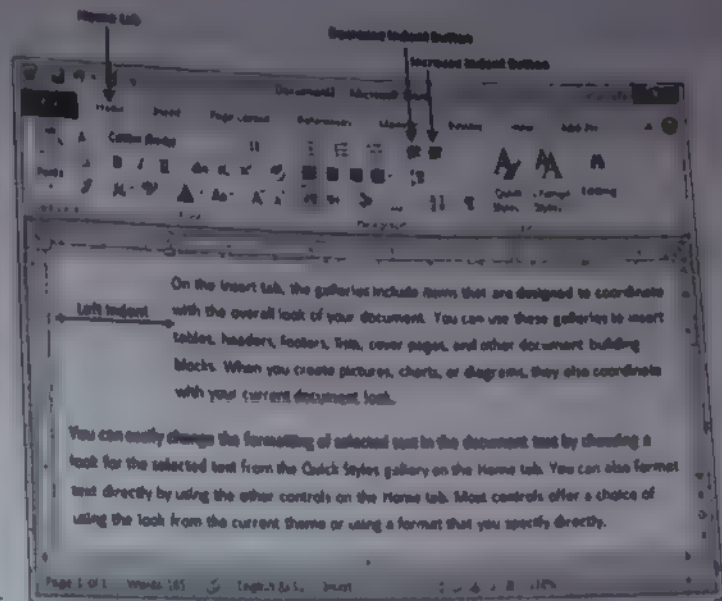


Figure 6.16 Indentation buttons

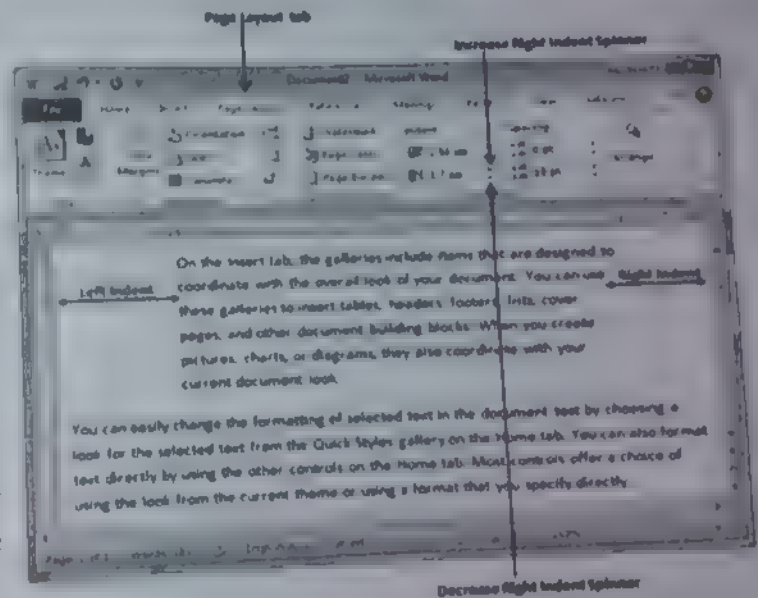


Figure 6.17 Indent Spinner button

6.4.2 Line and Paragraph Spacing

The vertical space between two lines can be adjusted in Word. While more space between two lines makes the text more readable and clear, less space between two lines, on the other hand, helps you to insert more text on a single page. Follow the steps given below to increase or decrease the line or paragraph spacing:

Step 1: Select the text.

Step 2: In the Paragraph group of Home tab, click on Line and Paragraph Spacing as shown in Fig. 6.18.

Step 3: Select the appropriate option. If you click on the Line Spacing option, the Paragraph dialog box (Fig. 6.19) will be opened, from where you can adjust the line spacing with more accuracy.

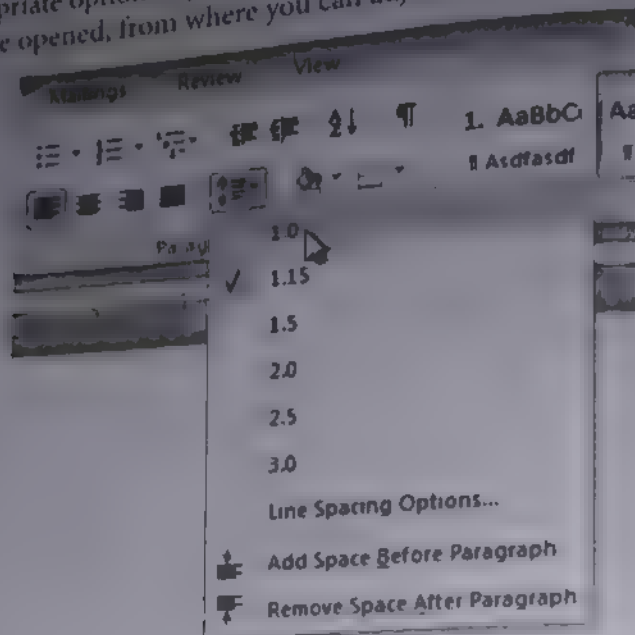


Figure 6.18 Line and Paragraph Spacing options

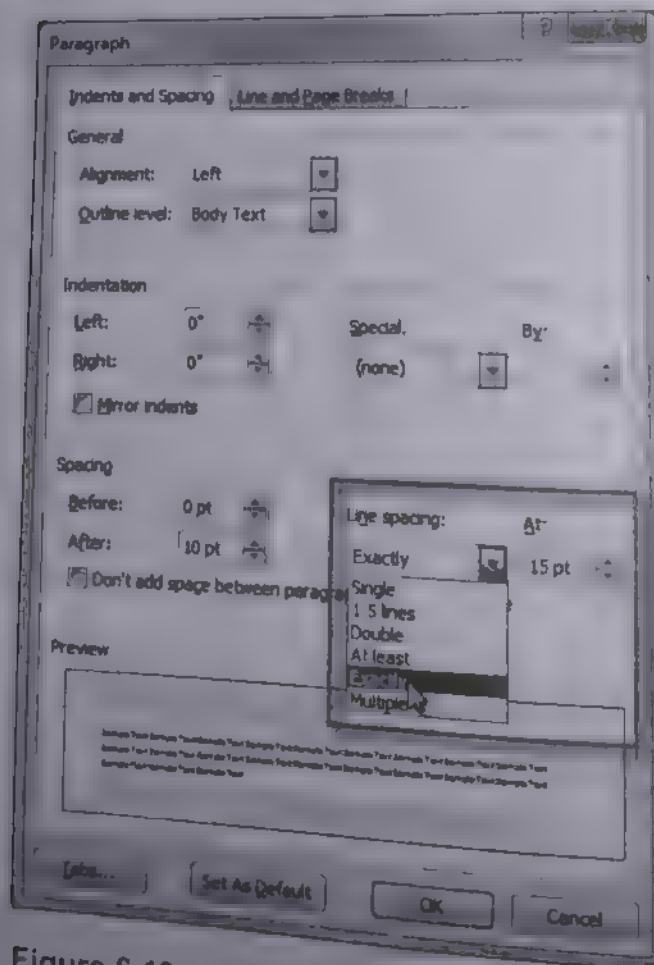


Figure 6.19 Paragraph dialog box



6.4.3 B

Microsoft W
leted list is c
or numbered

Step 1: Sel

Step 2: Cl
list

Step 3: Cl

Line spacing can be either measured in points or number of lines. It is usually measured in lines but if you select At least or Exactly in the Paragraph dialog box, it will be measured in points.

6.4.3 Bullets and Numbering

Microsoft Word provides bullets and numbers to maintain a list of items in a presentable order. A bulleted list is called unordered list and a numbered list is called ordered list. If you want to make a bulleted or numbered list, then follow the steps given below:

- Step 1:** Select the list of text.
Step 2: Click on the triangle next to the Bullet button to display a list of bullets. To create a numbered list, click on the triangle next to the Numbering button to display a list of numbers.
Step 3: Click on any of the bullet style to select it as shown in Fig. 6.20.

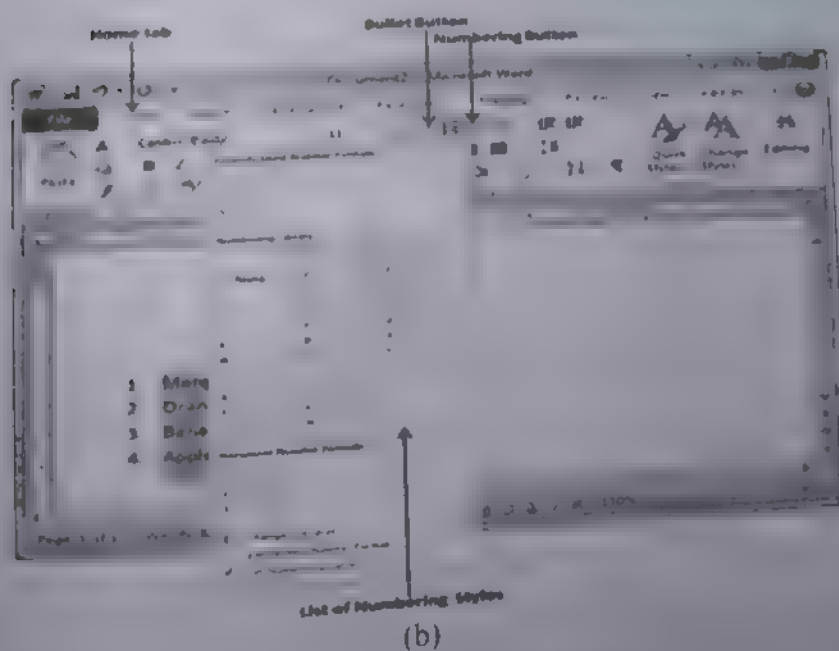
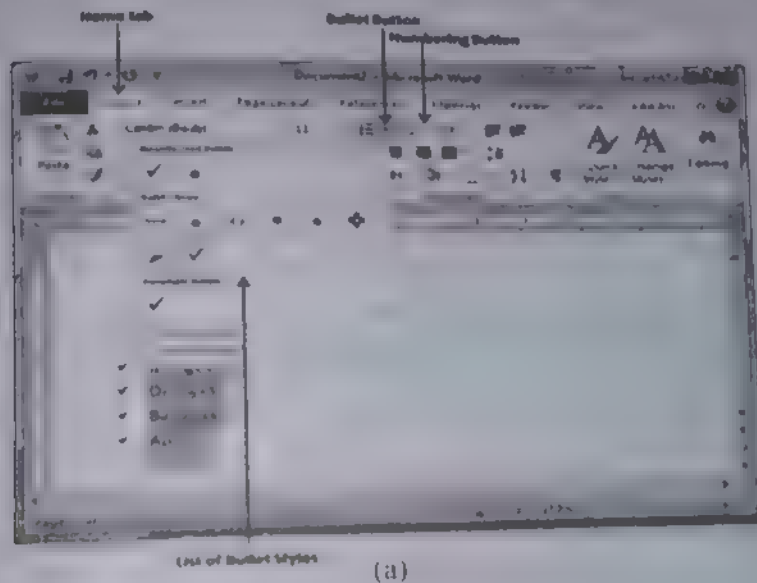


Figure 6.20 Adding bullets and numbers



An alternative way for indenting a paragraph, aligning text, changing line and paragraph spacing, and adding bullets and numbers to the text is to select the desired sentences or paragraphs in a document and then right click on them. Choose the appropriate option (i.e., Paragraph Bullets, and Numbering) from the list to perform the required task on the selected text. Based on the selected option, a dialog box opens (as shown in Figs 6.19 and 6.20) and then you can further perform your actions.

6.4.4 Case Changing

We use text in different cases to make it more readable. For example, we put headings in upper case and content under the heading in Sentence case in which only the first character is capital and rest of the characters is in small letters. Although you can capitalize character(s) by pressing the Caps Lock key or typing the character(s) while holding the Shift key, Word also provides features to change the case after the characters have been typed. Follow the steps given below to change the case of text:

Step 1: Select the text.

Step 2: In the Home tab, click on the triangle next to the Change Case button as shown in Fig. 6.21.

Step 3: Click on the option to select a case for your selected text.

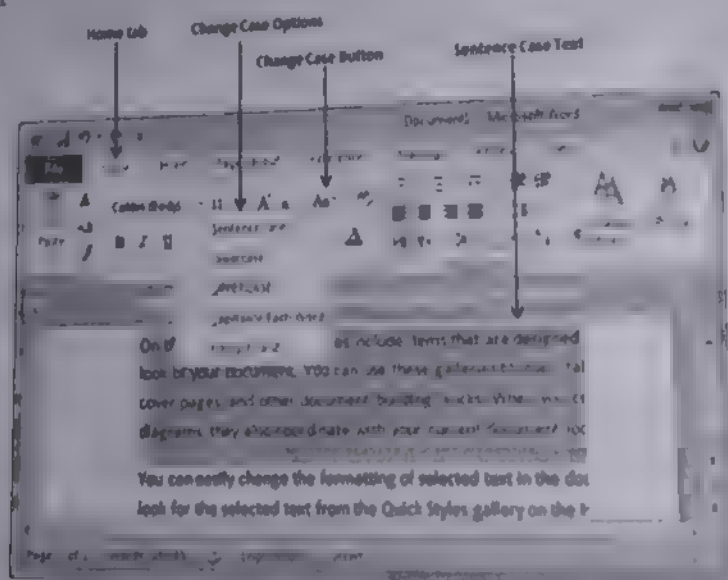


Figure 6.21 Changing case



In Capitalize Each Word option, every first character of every selected word is in capital. Toggle text will reverse the case of a text. It will make a lower case character an upper case character and vice versa.

6.4.5 Finding and Replacing Text

Word allows users to easily locate specific word(s) in the text. They can even replace those word(s) with (an)other word(s). The following are the steps to Find and Replace text in a Word document:

Step 1: Click on
Step 2: In the Find and Replace dialog box, enter the text you want to find in the Find what box. The Find operation will start from the beginning of the document.

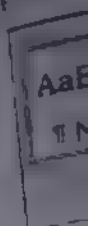


Fig.

Step 3: The Find and Replace dialog box will appear. In the Find what box, enter the text you want to find. In the Replace with box, enter the text you want to replace it with. Click on the Find Next button to find the next occurrence of the text. Click on the Replace button to replace the text.



6.4.6 Spelling and Grammar

In Word, the Spelling and Grammar checker will check the spelling and grammar of the text. It will highlight the words that are misspelled or have grammatical errors. You can click on the right-click menu to see the suggestions for the correct spelling or grammar.

Step 1:

Step 2:

Step 1: Click on the Home tab.

Step 2: In the Editing group, click on Find if you just need to find a text or click on Replace if you want to find operation to be followed by Replace as shown in Fig. 6.22

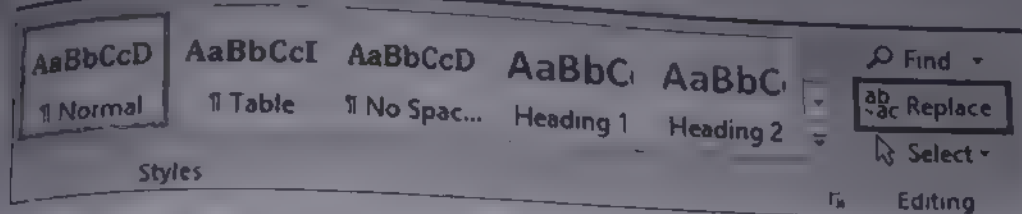


Figure 6.22 Replacing text

Step 3: The dialog box shown in Fig. 6.23 appears. In the Find what box, enter the text you want to search and then probably replace.

Step 4: In the Replace with box, enter the replacement text.

Step 5: Click on Replace. To replace all occurrences of the text in the document, click Replace All.

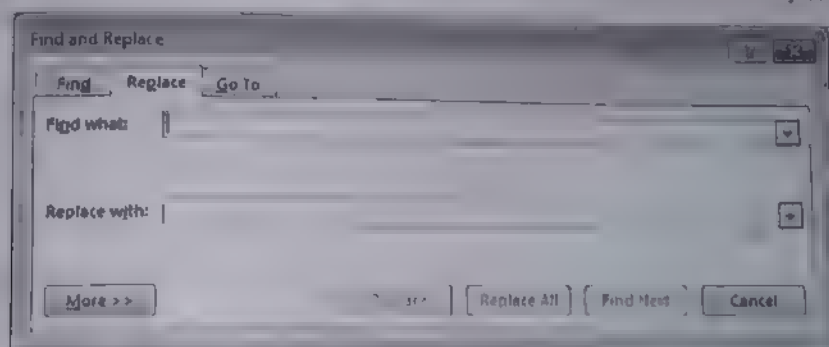


Figure 6.23 Find and Replace dialog box



To skip the current occurrence of the text and proceed to the next, click on Find Next.

6.4.6 Spell Check

In Word, users can check the spelling and grammar of the text all at once by running the spelling and grammar checker. The check can also be done automatically and then corrections can be made by the user. Moreover, the automatic spelling and grammar off feature can be turned off as and when required. The steps to do this are:

Step 1: Click the File tab to go to Backstage view, then click on Options as shown in Fig. 6.24.

Step 2: From the dialog box that appears, click on Proofing on the left side.

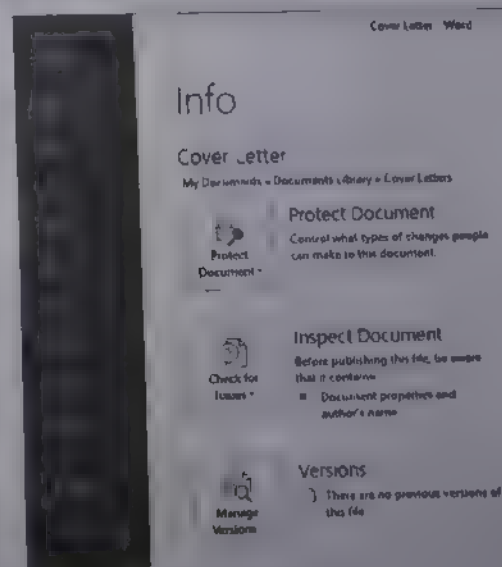


Figure 6.24 File tab

Step 3: On the right side, shown in Fig. 6.25, if you do not want Word to automatically check spelling, uncheck Check spelling as you type. If you do not want grammatical errors to be marked, uncheck Mark grammar errors as you type. If you do not want Word to check for contextual errors, uncheck Frequently confused words.

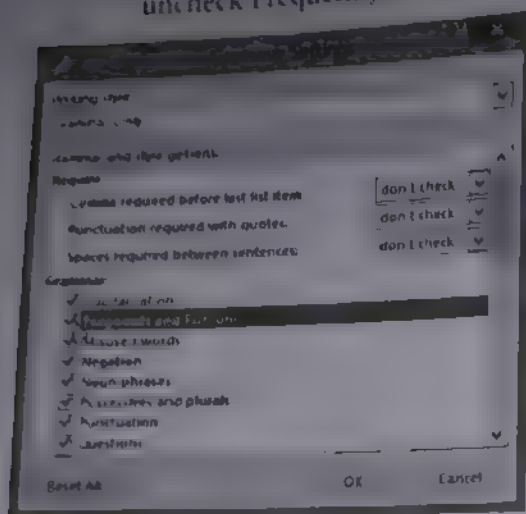


Figure 6.26 Grammar settings

6.4.7 Checking Grammar

By default, Word does not check for sentence fragments (incomplete sentences) and run-on sentences. To use this feature, turn it on by clicking on Settings. From the dialog box that appears as shown in Fig. 6.26, check the box next to Fragments and Run-ons

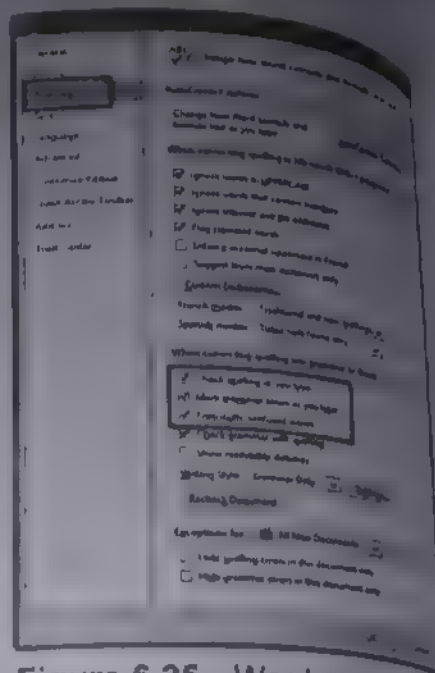


Figure 6.25 Word options

in Fig. 6.26, check the box next to Fragments and Run-ons

6.4.8 AutoCorrect

Word allows the AutoCorrect feature to correct typos, capitalization errors, and misspelled words. It is also used to automatically insert symbols and other pieces of text. By default, AutoCorrect has a list of typical misspellings and symbols, but users can modify it as per the need. Follow the steps given below to use the AutoCorrect feature.

Step 1: Click on File tab.

Step 2: Click on Options.

Step 3: Click on the Proofing tab and then click on AutoCorrect Options, as shown in Fig. 6.27.

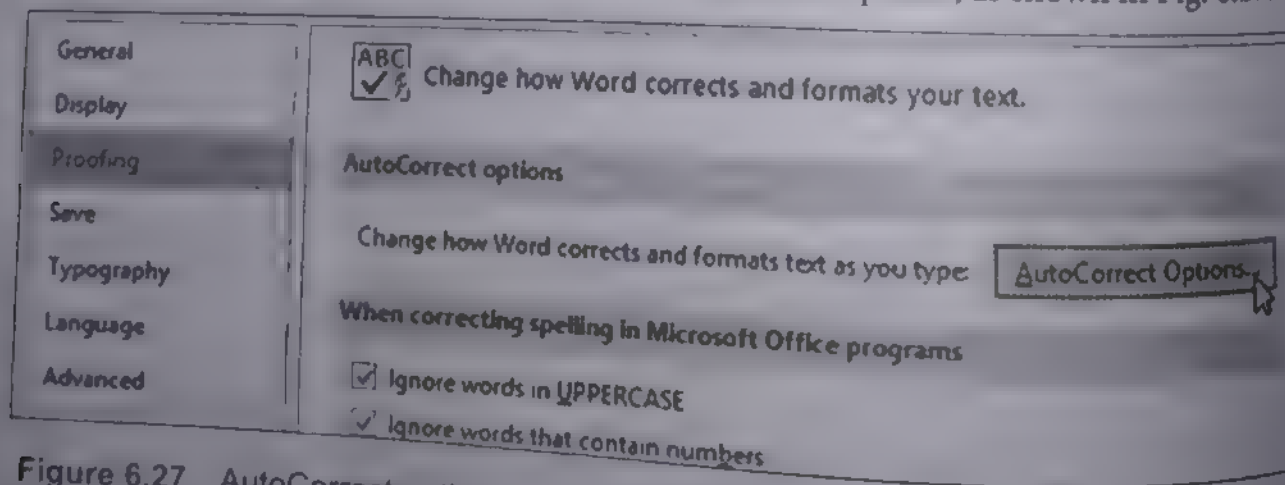


Figure 6.27 AutoCorrect options

You can either add or delete an entry in the list. To delete an entry, select the entry and click on Delete button. To add an entry, use the Replace With box (as shown in Fig. 6.28).

6.4.9

Users

Word

Step

Step

St

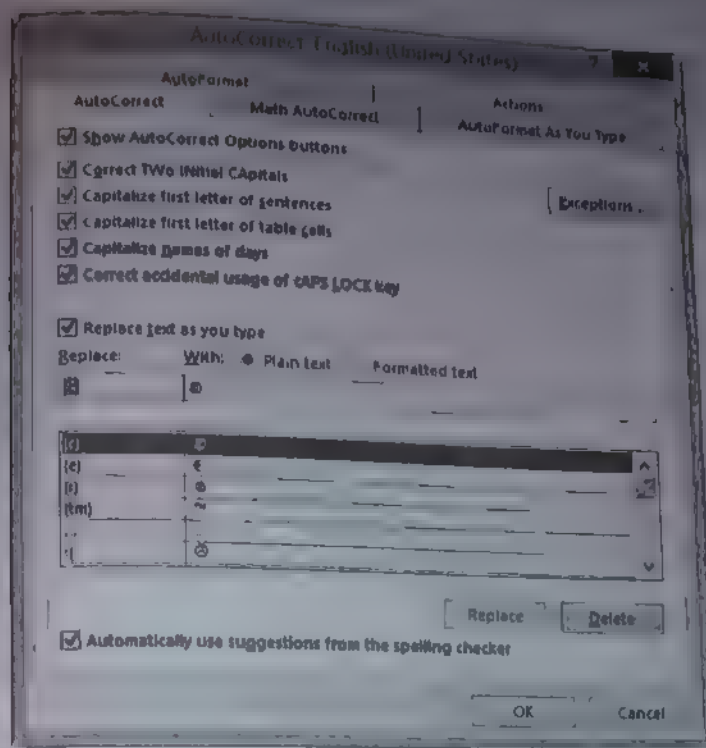


Figure 6.28 AutoCorrect options

6.4.9 AutoText

Users can insert text automatically in Word with blocks of preformatted text from the AutoText gallery. Word also inserts text automatically when the user has typed only a few characters.

Step 1: Select the text or graphics from a document that you want to add in the AutoText entry.

Step 2: Press Alt+F3 to open Create New Building Block box (Fig. 6.29), in which you should do the following.

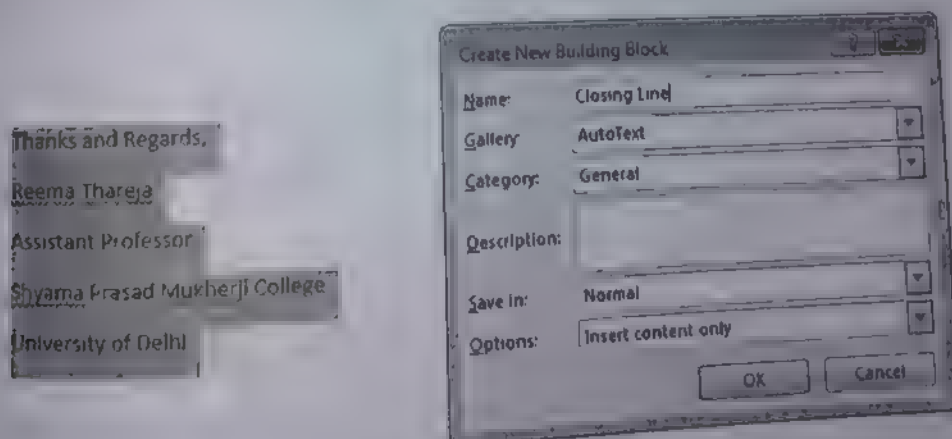


Figure 6.29 Creating AutoText

Step 3: Type a name of new AutoText. Make sure that from the Gallery drop down menu, AutoText is selected.

Step 4: Click OK to save the entry.

Now, to insert an AutoText entry in your text, follow these steps (Fig. 6.30):

Step 1: Click at the position where you want the text to be entered.

Step 2: Click on the Insert tab and then select the Quick Parts button in the Text group.

Step 3: Click on AutoText and select the Auto Text (by name) to be inserted.

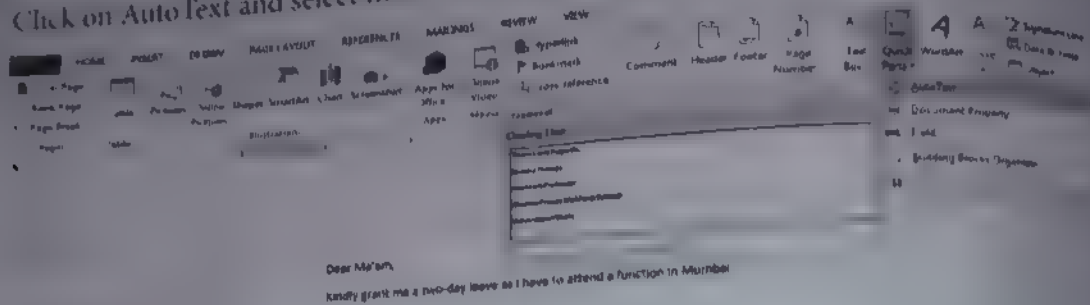


Figure 6.30 Inserting AutoText

6.5 HEADERS AND FOOTERS

Headers and/or footers are added to include important information about the document. This information may include the document's title, date of creation, page number, etc. Headers and footers are mainly used to display recurring information (like page number, name of the chapter, author's name, etc.) on the pages to help the reader keep track of your document's content.

Headers and footers appear on every page. Before using them, make sure they don't make your page look cluttered. The default font size is usually small enough to minimize distraction, but you can always change the font size and placement.

There is no general rule as to what should come in the header and what should come in the footer. Follow the steps given below to insert header and footer in your presentation.

Headers and footers can help keep longer documents organized and make them easier to read. The text entered in the header/footer will appear on each page of the document.

To insert a header/footer, follow these steps:

Step 1: Click on the Insert tab.

Step 2: Click either the Header or the Footer command.

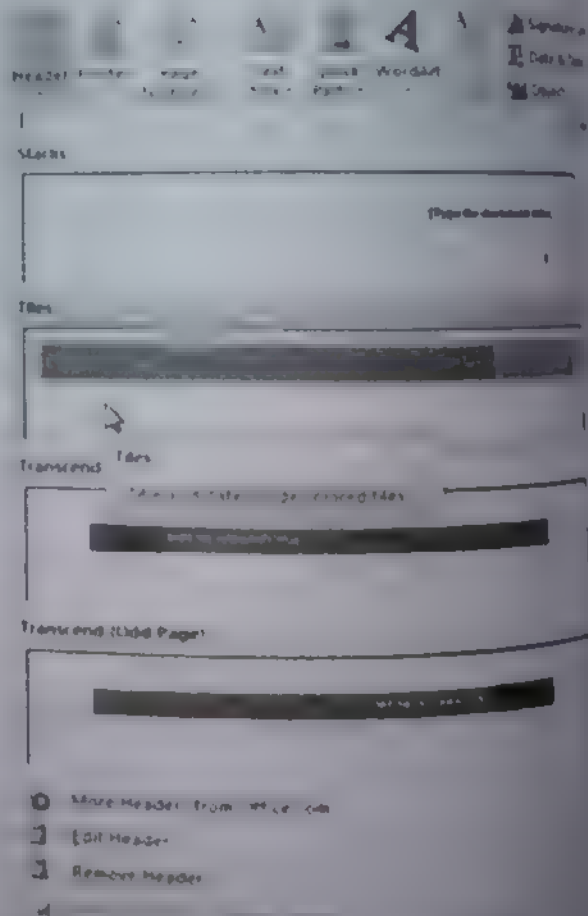


Figure 6.31 Adding Headers

Step 3: From 6.31)

Step 4: You in the the

Step 5: Typ foot Dat in h wa can sec

Step 6: On



6.6 MA

Macros is automatic In Word effort to be run se

Recon

Step 1:

Step 2:

Step 3:

Step 4:

Step

- Step 3: From the drop down menu that appears (refer Fig. 6.31) select any option of your choice.
- Step 4: You will now see that the header/footer has appeared in the document and a Design tab is now present in the Ribbon.
- Step 5: Type the desired information into the header or footer. You can enter date/time by clicking on the Date & Time command in the Design tab as shown in Fig. 6.32. Check the Update Automatically if you want it to be updated and click on OK. Similarly, you can also enter page numbers in the Header/Footer section of the document.
- Step 6: Once you are finished, press the Esc key.

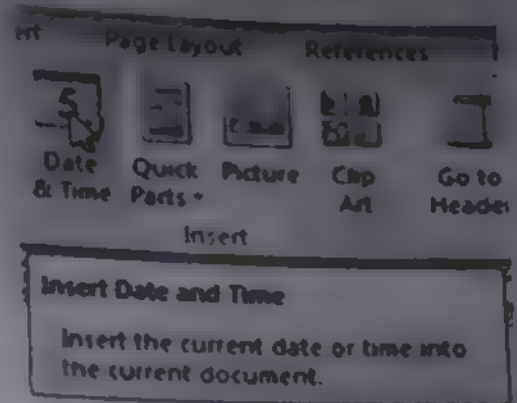


Figure 6.32 Date & time command



To edit header/footer double click on it and edit it.

6.6 MACROS

Macro is a series of commands that can be grouped together as a single command to accomplish a task automatically.

In Word, users can use macros to automate frequently used tasks. Macros of course save time and user's effort to do the same task repeatedly. To use a macro, users must first create and save a macro that can be run several times as and when required.

Recording a Macro with a Button

- Step 1: Click on the View tab.
- Step 2: Click on Macros from Macros group and then select Record Macro (Fig. 6.33).
- Step 3: In the dialog box that appears, shown in Fig. 6.34, type a name for the macro and then click on OK.

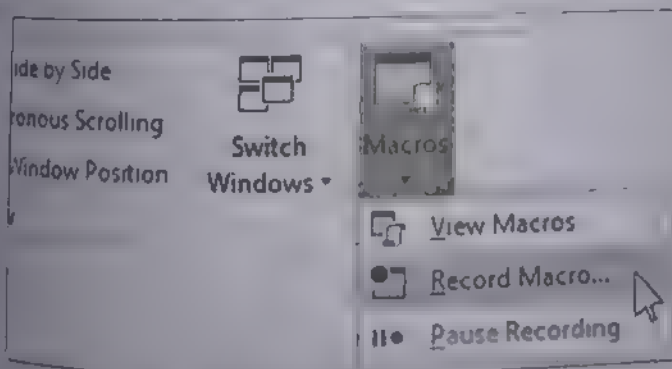


Figure 6.33 Record macro button

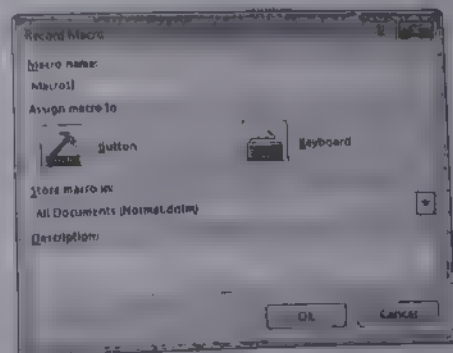


Figure 6.34 Recording a macro

- Step 4: The macro will start recording automatically. Just click the commands or press the keys for each step in the task.
- Step 5: To stop recording, click on View tab. From Macros select Stop Recording.

Running a Macro

Step 2: Click on Run.

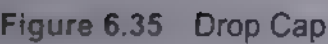
Step 2: Click on Klm.



6.7 DROP CAP

Have you ever noticed while reading books, novels, and newspapers that new chapters start with the first letter larger than the others? This is called Drop Caps which is especially used to draw the attention of reader and improve the appearance of the document. Follow these steps to use Drop Cap in your Word document.

Step 2: In the Text group, click on the Drop Cap button. From the menu that appears (Fig. 6.35), select the desired style.



Step 3: The selected drop cap will be automatically added to the document. You can then resize the drop cap to make it of the desirable size as shown in Fig. 6.36.



The Government of India enacted the Information Technology (IT) Act with an objective to

Figure 6.36 Resizing a drop cap letter

6.8 TABLE MANIPULATION

A table is a structure of vertical *columns* and horizontal *rows*. The intersection of a row and a column is called a *cell*. Each cell can store some data. This data can be a text, picture, or any other thing. Like ordinary text, you can also format a table to make it look attractive, easily readable, and understandable. Usually, the top row in a table is used for writing headings or for giving some informative instruction.

6.8.1 Drawing Table

To draw a table in Word, follow the steps given below:

Step 1: Click on the Insert tab.

Step 2: Click on Table to display a simple grid. As you move the mouse over the grid cells, a table with selected number of rows and columns is created in the document as shown in Fig. 6.37.

Step 3: Click on the square representing the lower-right corner of your table to finally create the table in your document.

Step 4: After creating the table, Word goes to table design mode where you get many options to work in the table.

Step 5: This is an optional step. If you want to have a fancy table, then click on Table Styles button (refer Fig. 6.38) to display a gallery of table styles. When the mouse is moved over any of the styles, the table is automatically updated in the document.

Step 6: Click on the desired style to apply the style to your table.

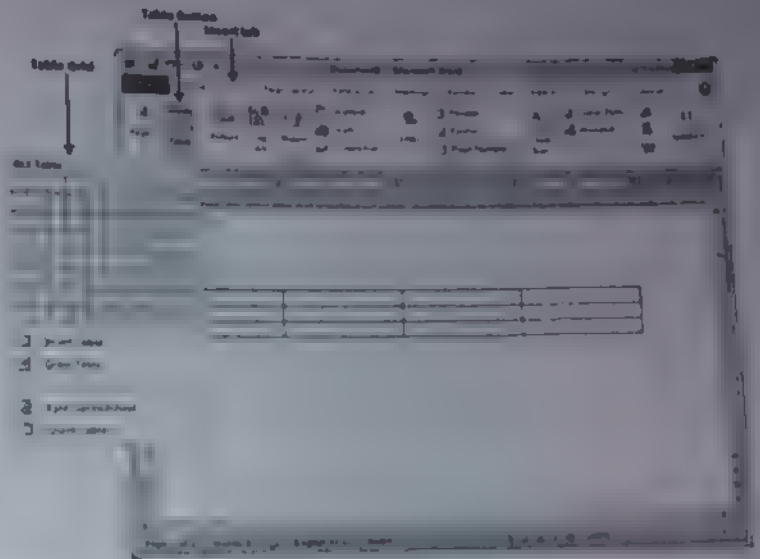


Figure 6.37 Inserting table

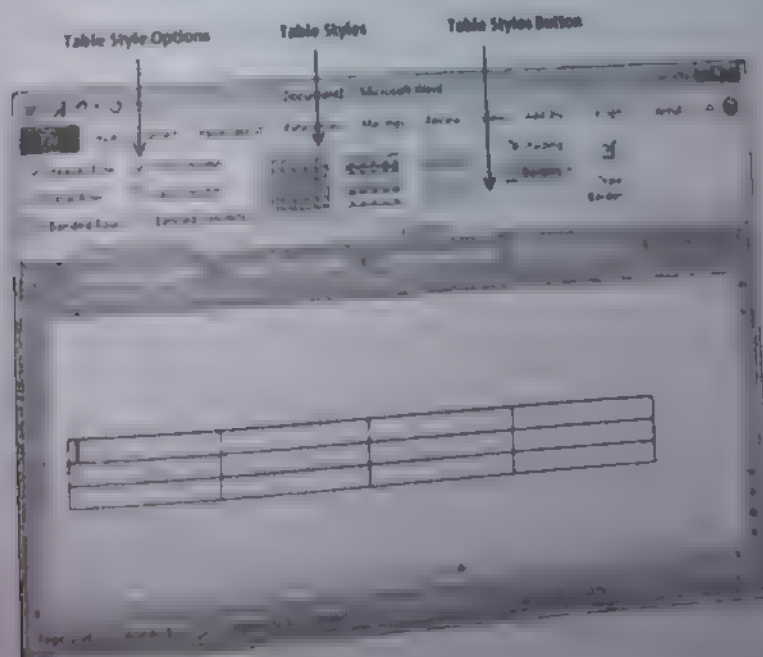


Figure 6.38 Table styles button

6.8.2 Changing Cell Width and Height

Changing the height and width of cells in a table is called *resizing* the table.

Step 1: Bring the mouse pointer inside the table and right click on it.

Step 2: Click on Table Properties.

Step 3: There are four tabs. Click on Row to increase or decrease its height as shown in Fig. 6.39(a). Click on Column to increase or decrease its width as shown in Fig. 6.39(b).

Step 4: Click on OK to apply the changes.

Alternative Way of Changing Column Width and Row Height

Another way to change column width is as follows:

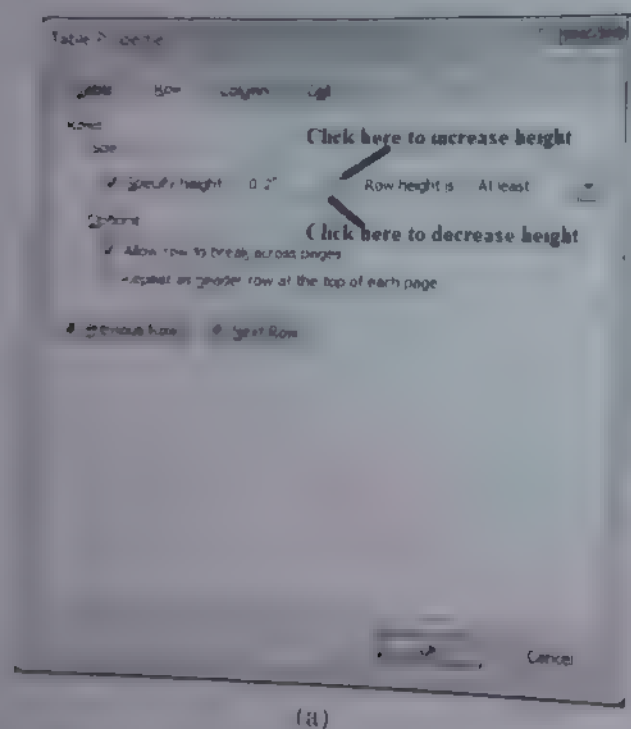
Step 1: Rest the cursor on right side of the column boundary until it becomes a resize cursor \leftrightarrow .

Step 2: Drag the boundary until the column has the desired width.

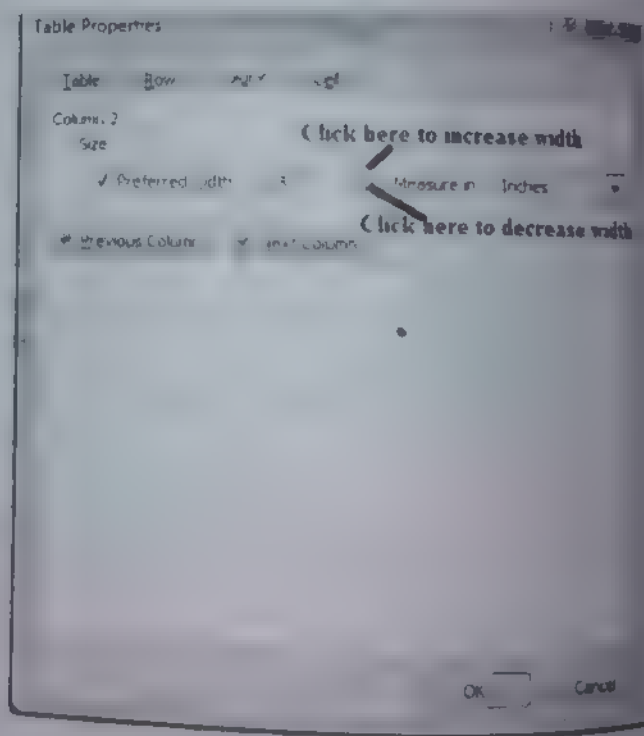
Follow the steps given below to make the columns in a table automatically fit the contents:

Step 1: Click on the table.

Step 2: Click on the *Layout* tab.



(a)



(b)


Figure 6.39 Row height and column width

Step 3: In the *Cell Size* group, click on *AutoFit*, and then click on *AutoFit Contents*.



You can even use the ruler to resize the column's width. Simply select a cell and drag the markers on the ruler. To see the exact measurement of the column on the ruler, keep the **Alt** key pressed while dragging the marker.

Another way to change row height is as follows:

- Step 1:** Rest the mouse pointer on the row boundary until it becomes a resize pointer 
- Step 2:** Drag the boundary until it reaches the desired height.

You can also use ruler to do the same task.

6.8.3 Alignment of Text in Cell

We can align text in the cell(s) of a table by following the steps given below:

- Step 1:** Select the cell(s).
- Step 2:** Right click on the cell(s).
- Step 3:** From the options, click on Cell Alignment as shown in Fig. 6.40.
- Step 4:** You will be given options to choose a particular alignment for your text. Click on the desired option as shown in Fig. 6.40.

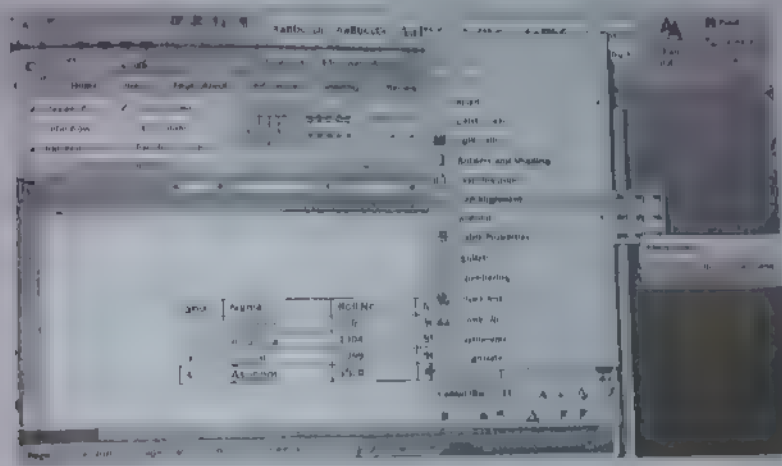


Figure 6.40 Cell alignment

6.8.4 Deleting/Inserting a Row/Column in a Table

Follow the steps given below to insert a new row in an existing table:

- Step 1:** Click the row where you want to add an additional row.
- Step 2:** Click on Layout tab as shown in Fig. 6.41. You will be given different options in the Rows and Columns group.
- Step 3:** Click on the appropriate option depending on where you want to insert the row.

Similarly, to delete a row from an existing table, perform the steps given below:

- Step 1:** Click the row that you want to delete from the table.
- Step 2:** Click on Layout tab.
- Step 3:** In the Rows and Columns group, click on the triangle next to the Delete button to see different options. Select Delete Rows.

Now to insert a column in an existing table, do the following steps:

- Step 1:** Select a column where you want to insert a new column.

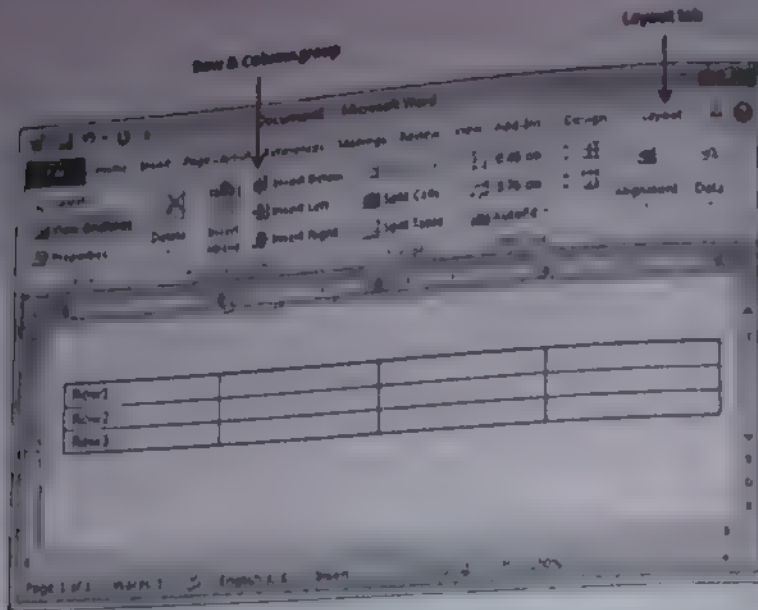


Figure 6.41 Table layout

Step 2: Click on the Layout tab.

Step 3: In the Rows and Columns group, click on Insert Left or Insert Right depending on whether you want to insert the new column to the left or right of the selected column as shown in Fig. 6.42.

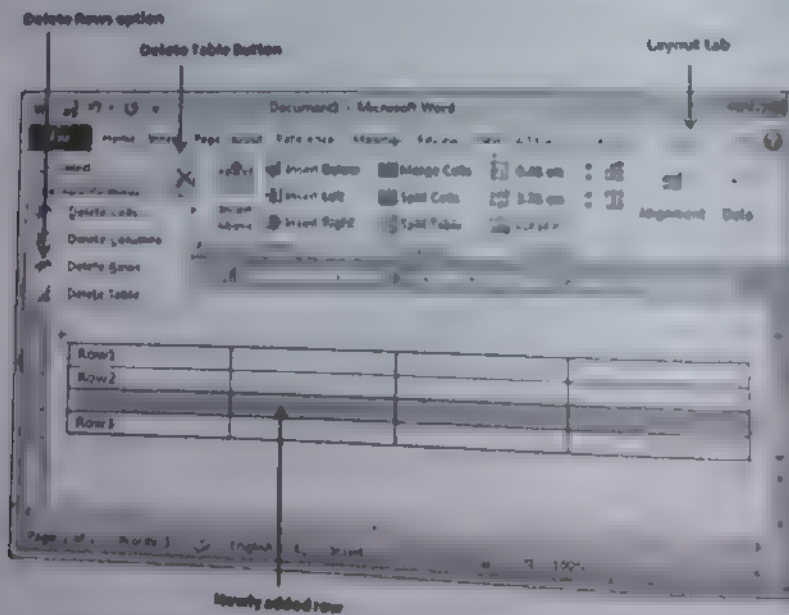


Figure 6.42 Delete rows/columns

Finally, to delete column(s) from a table, follow the given steps:

Step 1: Click on the column(s) to be deleted.

Step 2: Click on the Layout tab.

Step 3: In the Rows and Columns group, click on the triangle next to the Delete button to see different options. Select Delete Columns.

6.8.5 Borders

Microsoft Word allows you to add borders to a table.

To add borders to a table, follow the given steps:

Step 1: Select the table.

Step 2: Click on the Layout tab.

Step 3: Click on the Borders group.

Step 4: Click on the Borders button.

You can also add borders to a table by using the Borders button in the Borders group.

Step 1: Select the table.

Step 2: Click on the Borders button.

Step 3: Click on the Borders button.

Step 4: Click on the Borders button.

Step 5: Click on the Borders button.

Step 6: Click on the Borders button.

Step 7: Click on the Borders button.

Step 8: Click on the Borders button.

Step 9: Click on the Borders button.

Step 10: Click on the Borders button.

Step 11: Click on the Borders button.

Step 12: Click on the Borders button.

Step 13: Click on the Borders button.

Step 14: Click on the Borders button.

Step 15: Click on the Borders button.

Step 16: Click on the Borders button.

Step 17: Click on the Borders button.

Step 18: Click on the Borders button.

Step 19: Click on the Borders button.

Step 20: Click on the Borders button.

Step 21: Click on the Borders button.

Step 22: Click on the Borders button.

Step 23: Click on the Borders button.

Step 24: Click on the Borders button.

Step 25: Click on the Borders button.

Step 26: Click on the Borders button.

Step 27: Click on the Borders button.

Step 28: Click on the Borders button.

Step 29: Click on the Borders button.

Step 30: Click on the Borders button.

6.4.5 Borders and Shading

Microsoft Word allows you to apply borders on any or all of the four sides of a table. You can even add borders to its rows and columns to make the table more attractive.

To add borders in an existing table, follow the steps given below:

- Step 1: Select the table by clicking on the cross icon visible at the top-left corner of the table.
- Step 2: Click on the Border button in the Home tab (refer Fig. 6.43). A list of options will be displayed.
- Step 3: Click on an option to apply it to the table.
- Step 4: To remove border, click on No border option.

You can also apply borders by using the Borders and Shading dialog box.

- Step 1: Select the table by clicking on the Cross Icon.
- Step 2: Click on the Border button and select Border and Shading option as shown in Fig. 6.44.
- Step 3: In the Border and Shading dialog box, click on Border tab display to list border settings, styles, and colour.

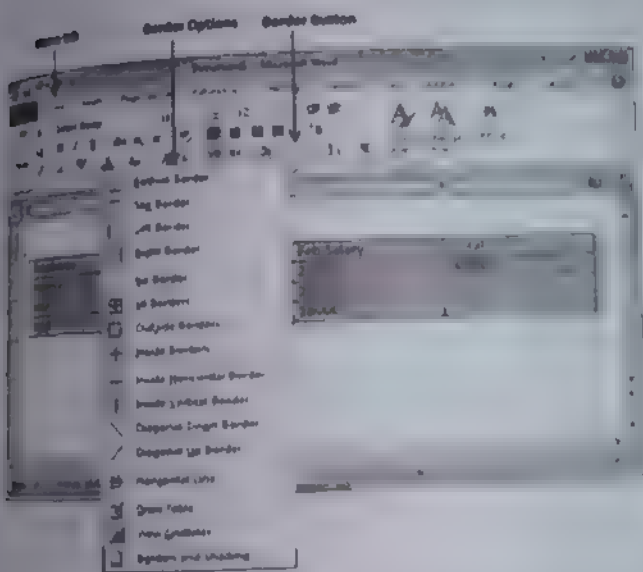


Figure 6.43 Border button

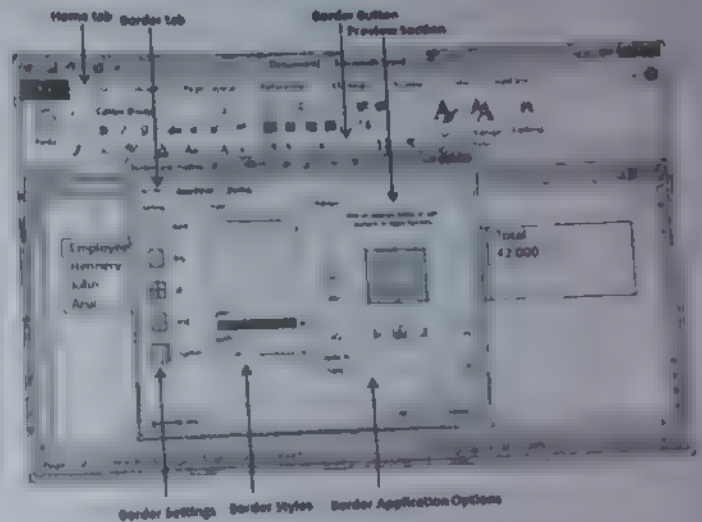


Figure 6.44 Borders

Step 4: Customize the border by selecting colour, width, and style for the border.

Now, to add shades on a selected table or its rows or columns follow the steps given below:

- Step 1: Select row(s) or column(s).
- Step 2: Click the Border button to display a list of options to put a border.
- Step 3: Select the Border and Shading option.
- Step 4: In the Border and Shading dialog box, click on the Shading tab (refer, Fig. 6.45) which will display options to select fill, colour, and style and whether this border should be applied to cell or table or selected text.
- Step 5: Make the settings and click on OK to apply the changes.



Figure 6.45 Shading

6.8.6 Merging Cells

Two or more cells in the same row or column can be merged or combined together to form a single cell. For example, you can merge cells horizontally to create a table heading. To merge cells, follow the steps given below:

Step 1: Select the cells that you want to merge.

Step 2: Right click and select Merge Cells from the menu as shown in Fig. 6.46.

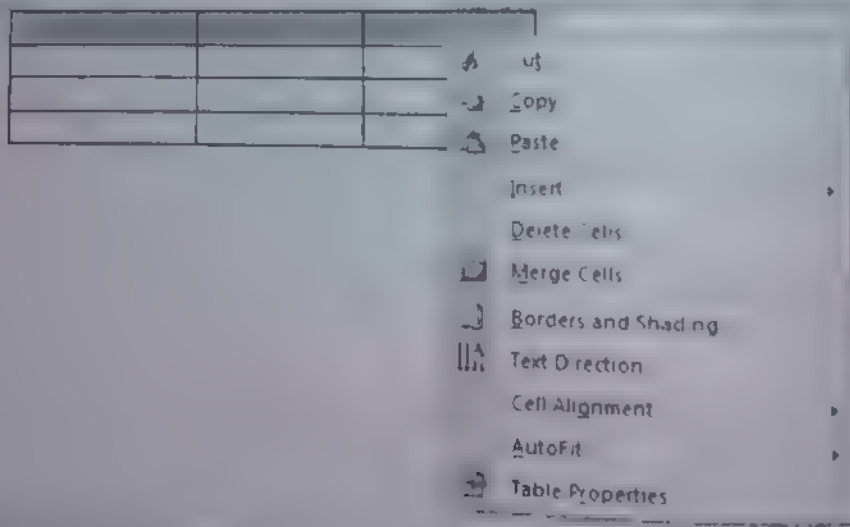


Figure 6.46 Merge cells

6.8.7 Splitting Cells

You can split a cell into two or more cells by performing the following steps:

Step 1: Right click in a cell.

Step 2: From the menu, select Split cells option as shown in Fig. 6.47.

Step 3:



6.8.8

To Adjust
of the
handle.

To Adjust
resize

To Adjust
You v

6.8.

Then
margin

Step

Step

Step

Step

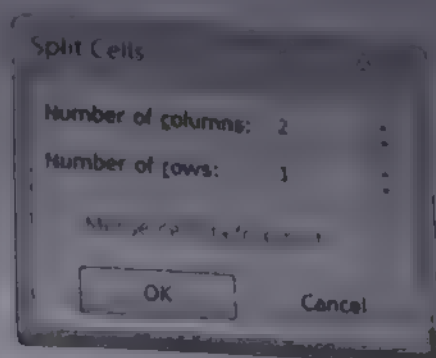


Figure 6.47 Split cells

Step 3: Enter the number of rows and columns in which you want to split the cell.



You can also merge/split cells from the Layout tab by clicking on the desired option in the Merge group.

6.8.8 Adjusting Cell Size by Dragging

To Adjust the Table Size In Print Layout view, move the mouse over the table. At the lower right corner of the table, you will get a resize handle as shown in Fig. 6.48. The moment you rest the mouse on this handle, it becomes a double-ended arrow. Drag the arrow until the size of the table becomes what you want.



Figure 6.48 Adjusting cell size

To Adjust the Column Size In Print Layout view, click on the boundary of the column that you want to resize. You will get a plus-like handle. Drag it until the column's width is what you desire.

To Adjust the Row Size In Print Layout view, click on the boundary of the row that you want to resize. You will get a handle. Drag it until the row's height is what you desire.

6.8.9 Cell Margins

There is an alternate way of adjusting cell margins. Follow the steps given below to modify the cell margins of a table:

Step 1: Click anywhere on the table.

Step 2: Click on the Layout tab.

Step 3: From the Alignment group, click on Cell Margins as shown in Fig. 6.49.

Step 4: The Table Options dialog box will open. From the dialog box, you can make the following adjustments.

- Under Default cell margins, enter the measurement for the Top, Bottom, Left, or Right margins.

- Under Default cell spacing, select Allow spacing between cells check box, and then enter the desired measurement

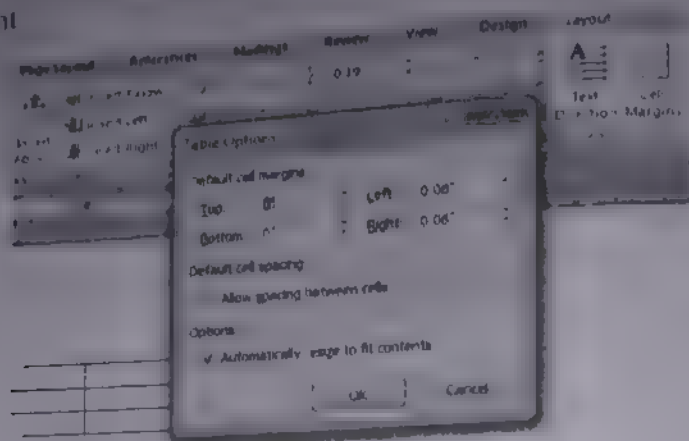


Figure 6.49 Cell margins command

6.8.10 Changing the Direction of Text in Table

Follow the steps given below to change the direction of text in one or more table cells:

Step 1: Select the cells for which you want to change the text direction.

Step 2: Click on the Layout tab.

Step 3: From the Alignment group, click on Text Direction as shown in Fig. 6.50.

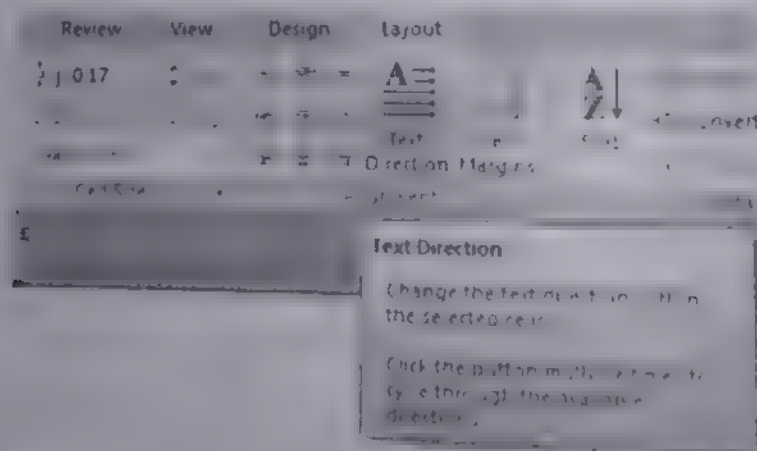


Figure 6.50 Text direction

Each time you click on Text Direction, a different direction is applied. Keep clicking the button until the desired direction is applied.

6.9 INSERTING PICTURES

Pictures are added in a document to illustrate important information or to make it look attractive. Pictures can be inserted (or copied) into a Word (or any Office) document either from the user's computer or from the Internet. Newer versions of MS Office no longer support clip art gallery. Follow the steps given here to insert pictures from your computer.

- Step 1: Click at the position in the document where you want to insert a picture.
- Step 2: Click on the Insert tab and then click Pictures as shown in Fig. 6.51.

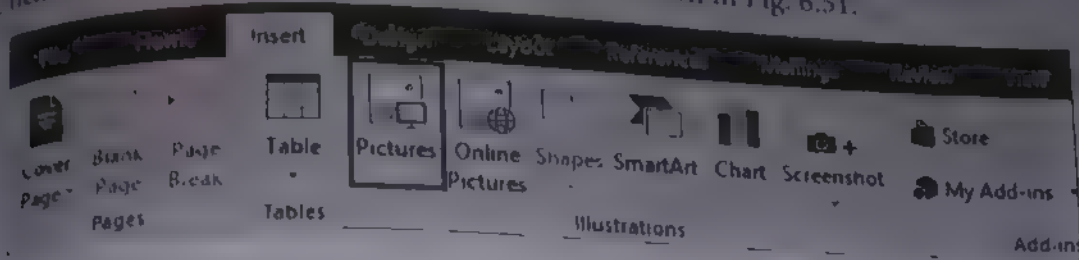


Figure 6.51 Insert picture

- Step 3: From the dialog box that appears, browse your computer to locate the picture you want to insert, select it, and then click on Insert.
- The picture will be inserted in the document. You can resize this picture to make it of desirable size.
- If the desired picture is not present in the computer, you can search for it online and then add it to the document. Follow these steps to add an online picture.
- Step 1: Click at the position in the document where you want to insert a picture.
- Step 2: Click on the Insert tab and then click Online Pictures as shown in Fig. 6.52.

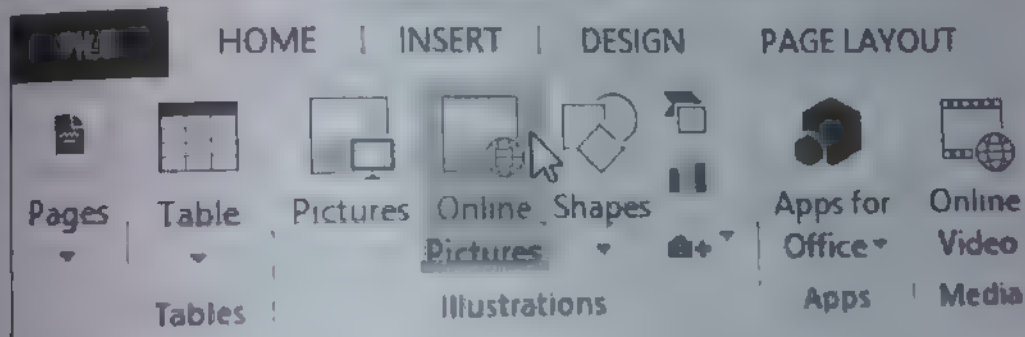


Figure 6.52 Insert an online picture

- Step 3: In the Bing Image Search, type the keyword to search for the picture and press the Enter key.
- Step 4: A dialog box showing search results will appear as shown in Fig. 6.53. Select the desired image, and click on Insert.

The image will appear in the document.

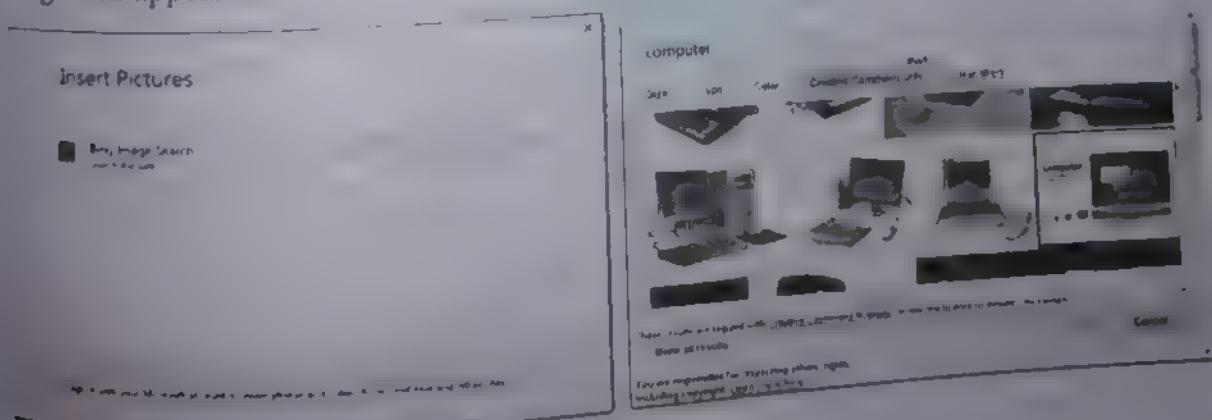


Figure 6.53 Online image search



Use the same steps to insert a video. Click on Insert tab followed by Online video, search for the video using Bing search, YouTube, or any other website and then insert it.

Changing Text Wrapping Settings

Once an image is inserted, it is very difficult to move it exactly where you want. This is because by default the image is placed in line with the text. To move the picture freely, you can use the **text wrapping** feature of MS Word. Follow the steps given below to wrap text around an image.

Step 1: Right click on the image around which you want to wrap text.

Step 2: From the pop up menu that appears, hover your mouse over the Wrap Text option.

Step 3: Select the desired option from another pop-up menu that appears (Fig. 6.54).

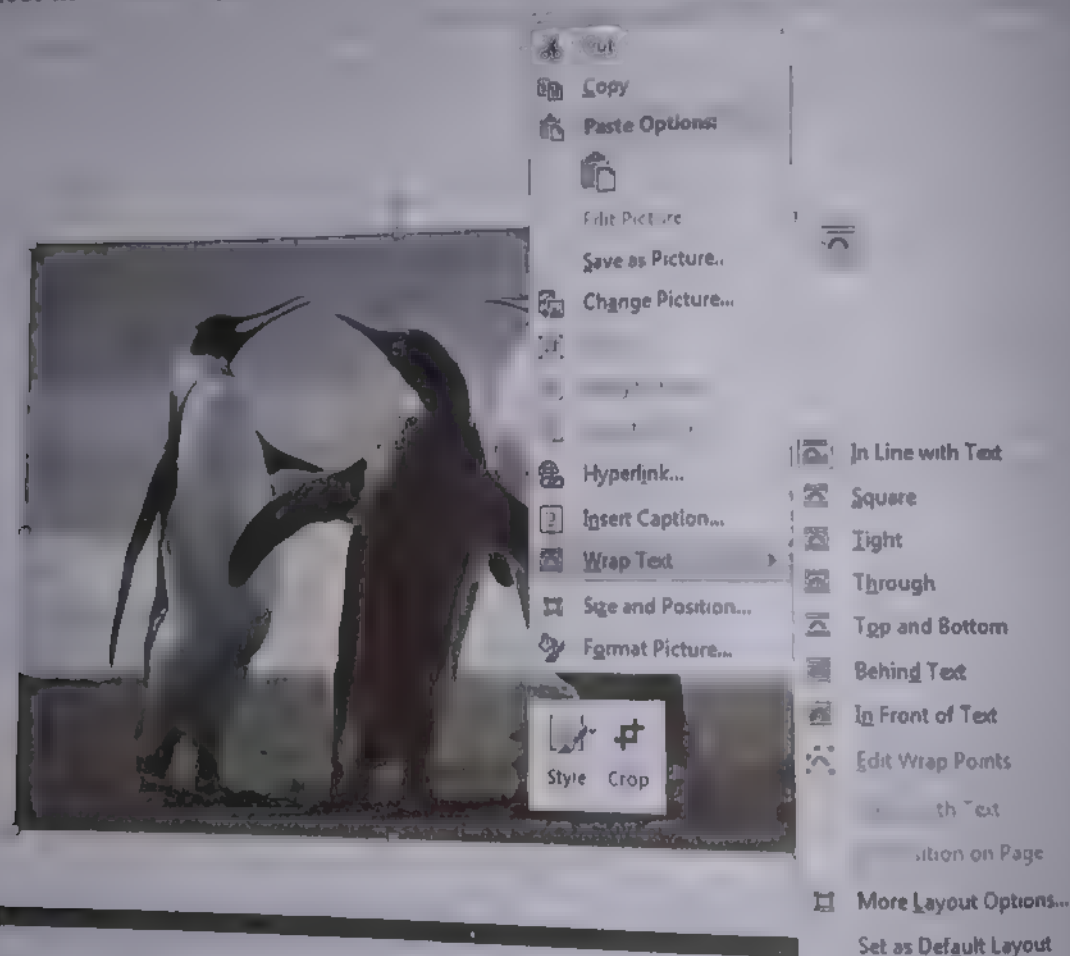


Figure 6.54 Text wrapping an image

6.10 INSERTING SHAPES

You can add a variety of shapes to your document. You can even add text in these shapes. These shapes also help you to draw images or other complex shapes. Some shapes already available include lines, basic geometric shapes, arrows, equation shapes, flowchart shapes, stars, banners, and call outs. Follow the steps given below to insert a shape in your document.

Step 1: Click on the Insert tab.

ch for

default
feature

- Step 1: Click on Shapes as shown in Fig. 6.55.
- Step 2: Select a shape from the drop-down menu.
- Step 3: Click and drag the mouse until the shape is of the desired size.
- Step 4: Release the mouse button. The desired shape of the selected size would be inserted in the document.

Once the shape is inserted, you can change its style by following the steps given below:

- Step 1: Select the shape. The Format tab can be seen in the Ribbon, click it.
- Step 2: Click the drop down arrow in the Shape Styles group to see more options.
- Step 3: Move your cursor over the styles (refer Fig. 6.56) to see a preview of the style.

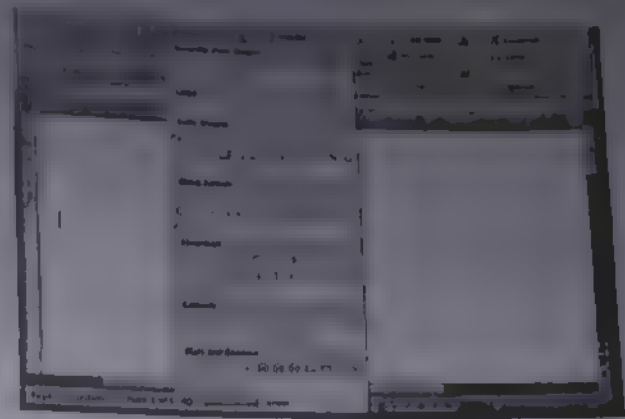


Figure 6.55 Inserting shapes

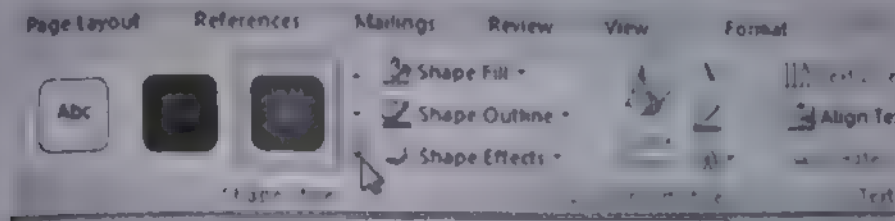


Figure 6.56 Shape styles

Step 4: Click on the desired style.

To change the colour of the shape, follow the steps given below.

- Step 1: Select the shape. The Format tab can be seen in the Ribbon, click it.
- Step 2: From the Shape Styles group, click on the upside down arrow next to the Shape Fill command (as shown in Fig. 6.57) to display a drop-down list.
- Step 3: Select the desired colour from the list.

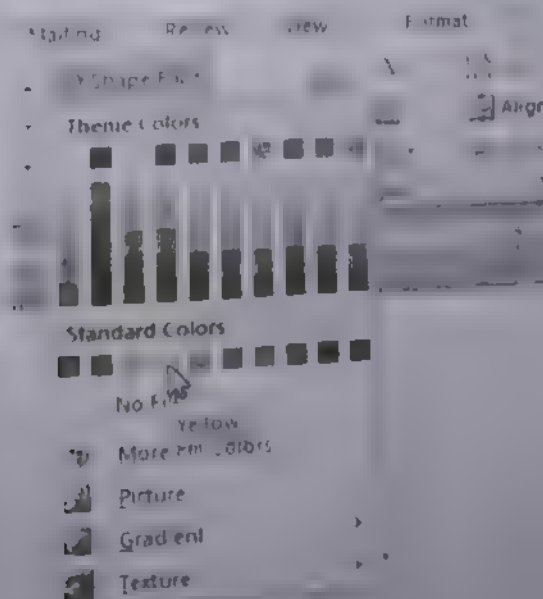


Figure 6.57 Shape fill



To change the shape's outline, click on Shape Outline in the Shape Styles group of Format tab. From here, you can change the colour, weight, and pattern of the line making the outline.

To add some effects to the shape, follow the steps given below:

Step 1: Select the shape. The Format tab can be seen in the Ribbon, click it.

Step 2: From the Shape Styles group, click on Shape Effects as shown in Fig. 6.58. A drop-down menu will appear.

Step 3: Move your mouse across the options to see a preview.

Step 4: Click on the desired shadow effect.

Some more interesting things that you can do to a shape are as follows:

Resize Select the shape. You will get resizing handles at each corner. Drag along each edge of the shape's bounding box to resize it.

Rotate Select the shape. You will get a green rotation handle at the top. When you hover the mouse over that handle, the cursor will change to a rotation symbol. Click and drag to the left or right to rotate the shape anticlockwise or clockwise respectively.

Delete Select the shape and press the Delete button on your keyboard.

Add Text Right click on Shape and select the Add Text option as shown in Fig. 6.59.

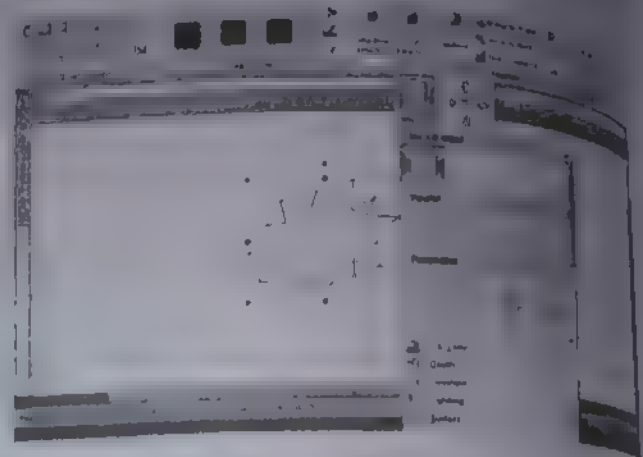


Figure 6.58 Shape effects

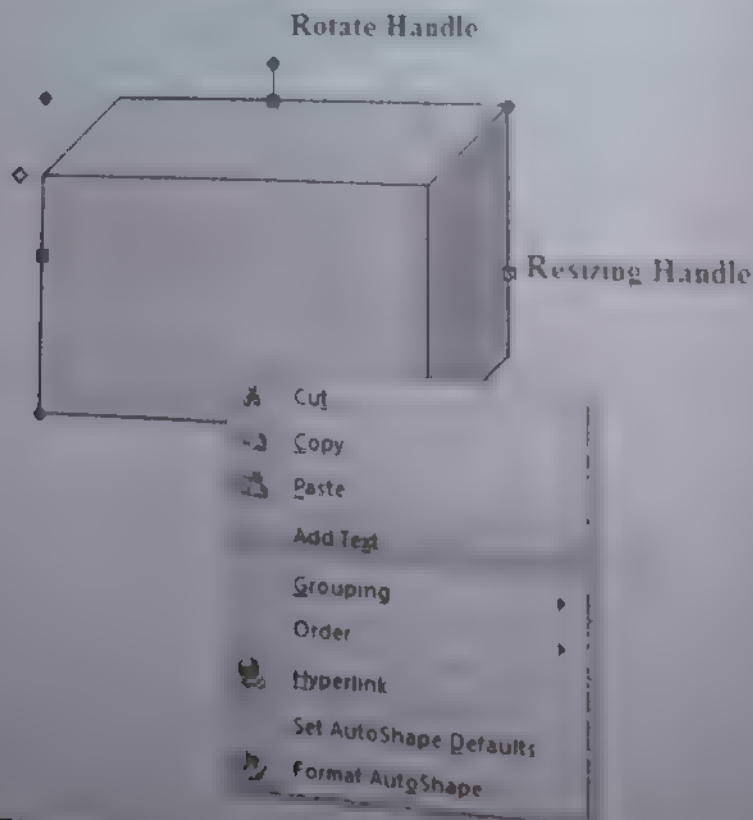


Figure 6.59 Adding text to a shape

6.11 INSERTING TEXT BOX

As other shapes, you can also insert a text box into a document to draw attention to specific text or have the ability to easily move text within a document. Since a text box is also a shape, you can add the same types of effects to them and even change their shape, resize them, move them, and modify the colour and outline. Follow the steps given below to insert a text box into a document:

- Step 1: Click on the Insert tab.
- Step 2: From the Text group, click on Text Box.
- Step 3: You can either select a text box (refer Fig. 6.60) from the drop down menu or click on Draw Text Box.
- Step 4: Click and drag on the document to create the text box.
- Step 5: Type your text inside the text box.

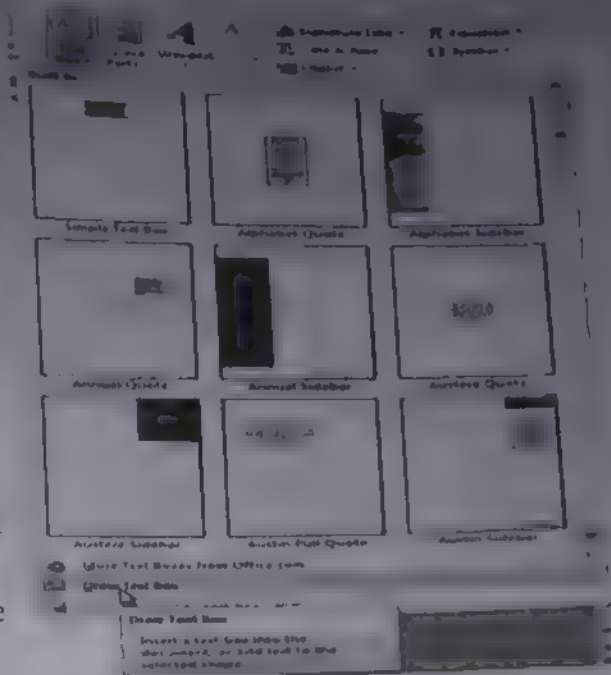


Figure 6.60 Inserting text box



To format a text box, you can resize it, change its shape, style, and colour, and apply effects as you do in case of any shape.

6.12 MAIL MERGE

Just like other types of shapes, you can also apply 3D effects to text boxes. You can give a 3D appearance to your text box or any other shape by using 3-D Rotation and Bevel. While 3-D Rotation gives the appearance that you are viewing an object from a different angle, Bevel adds thickness and a rounded edge to shapes. Follow the steps given below to apply 3D effect to any shape including text box.

- Step 1: Select the shape.
- Step 2: Click on the Format tab.
- Step 3: From the Shape Styles group, click on Shape Effects.
- Step 4: Move your mouse over 3-D Rotation. From the drop-down menu select one of your choice as shown in Fig. 6.61.

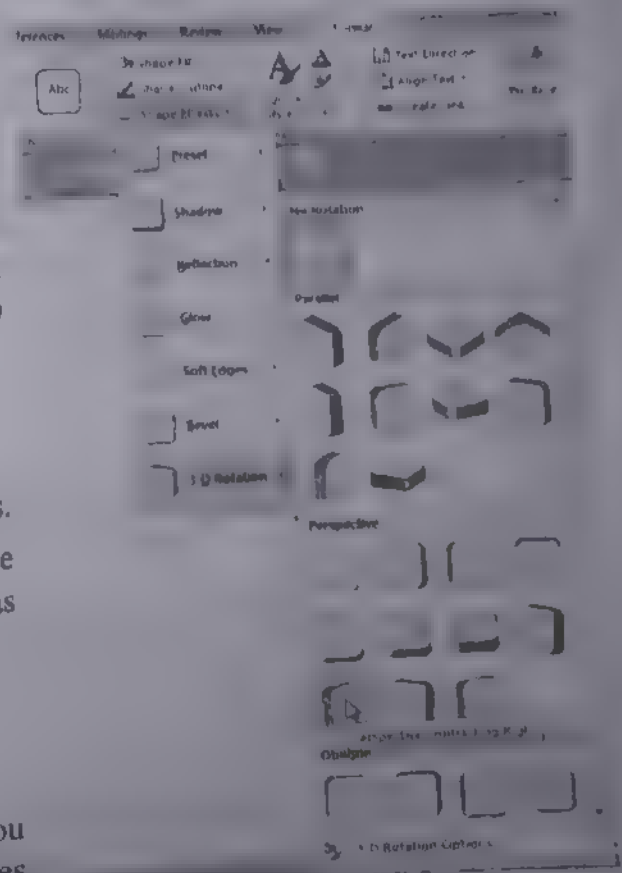


Figure 6.61 3D rotation

Mail merge is a feature in Microsoft Word that allows you to type in a single document and add a list of names besides other relevant information. It then combines or merges all

these things. Mail merge is especially used in cases where a single document has to be sent to multiple receivers.

For example, you have an important meeting in your office and you are asked to invite 100 delegates. Then instead of typing 100 separate documents, you just have to type a single document. Just provide the name and address of those 100 people, Word will automatically create 100 invitations and insert all their names so that you can take a printout and send them.

To use the Mail Merge feature of Word, follow the steps given below:

Step 1: Open MS Word. Click on Mailings tab and then on Step-by-Step Mail Merge as shown in Fig. 6.62.

Step 2: On the right side of the screen, a list of options is displayed regarding the type of document. Select Letters and click on Next shown in the bottom of the column.

Step 3: Another list of options (shown in Fig. 6.63) is displayed to confirm the starting document. Select Use the current document and click on Next.

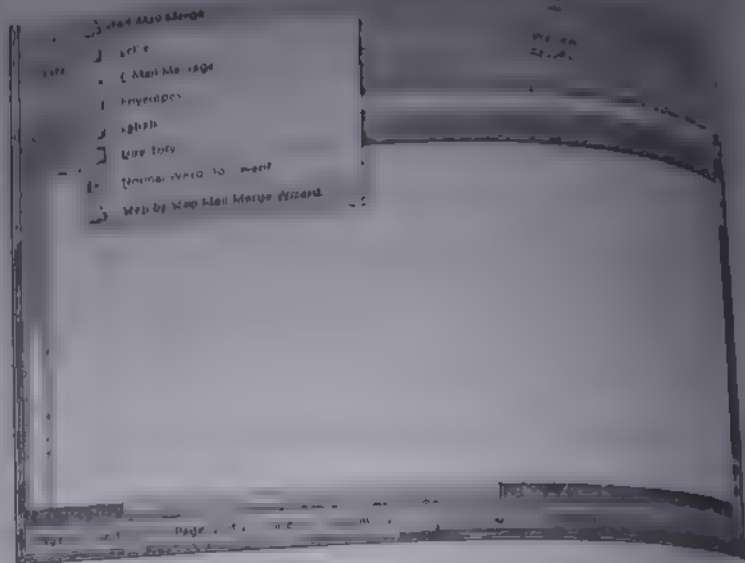


Figure 6.62 Mail merge

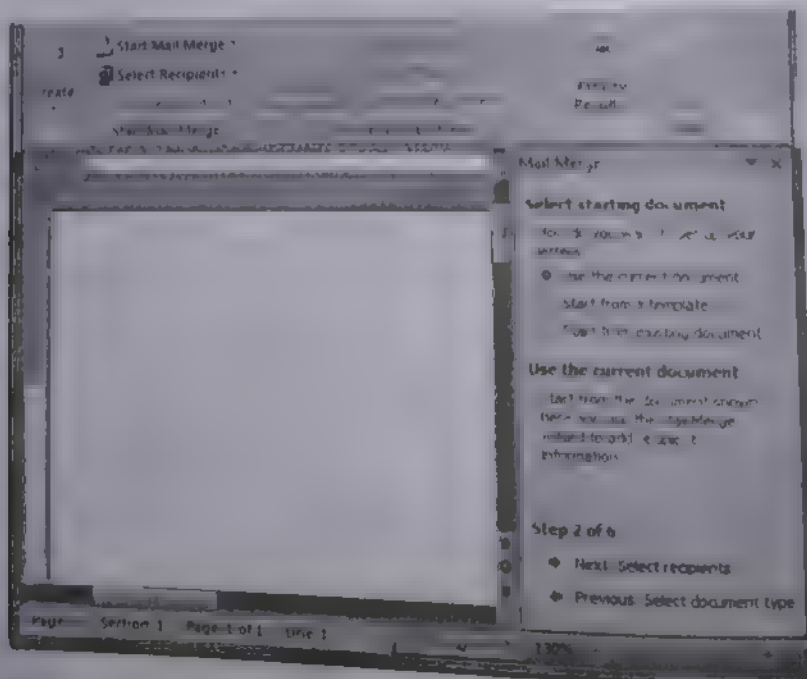


Figure 6.63 Options

Step 4: An option box is again displayed to allow you to choose a list of recipients. Select Type a New List then you will get an option to Create. Click on it and enter the details of the recipient (Fig. 6.64), and save this information at any location in your computer. After saving information click on Next.

Step 5: Click on Address Block. Unselect the Insert the recipient's name (Fig. 6.65) and click OK. Then press Enter and click on Greeting Line. Click OK in the dialog box. Press Enter and then type the letter. Now click on Next.

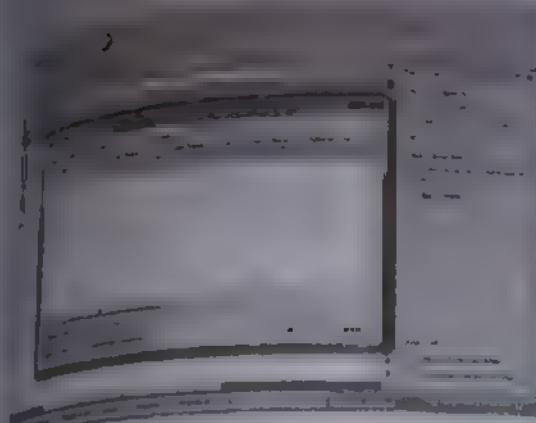


Figure 6.64 Create existing list

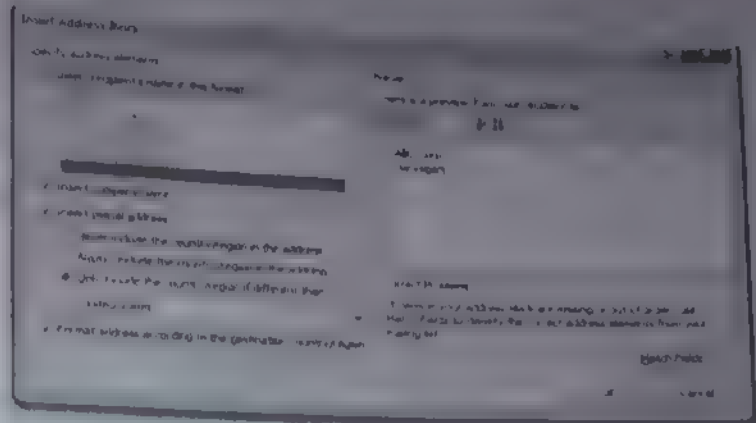


Figure 6.65 Address block

Step 6: The letter is ready for the first recipient, to check the letters for other recipients, click on >>, in the right side of the Option bar.

Step 7: Click on Complete the merge. Now you can either edit individual letters or print all of them as shown in Fig. 6.66.

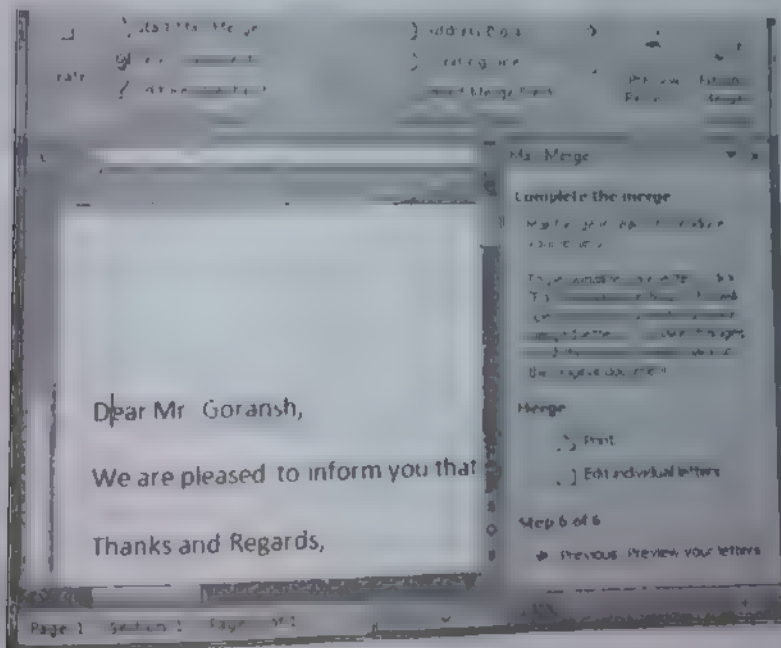


Figure 6.66 Print or edit

6.12.1 Creating Envelopes and Labels

Step 1: Open MS Word and click on Mailings.

Step 2: Click on Labels. A box is shown; fill in the details as shown in Fig. 6.67. You can now print the label either by having one label on a single page or having multiple labels on a page. For envelopes, go to Mailings and select Envelopes. The box that pops up asks you to enter the delivery and return address. Fill in the details and click Print as shown in Fig. 6.68. To change the position

where the return or recipient's address is printed, click on Feed which is next to Preview. You can also add graphics on the label or envelope.

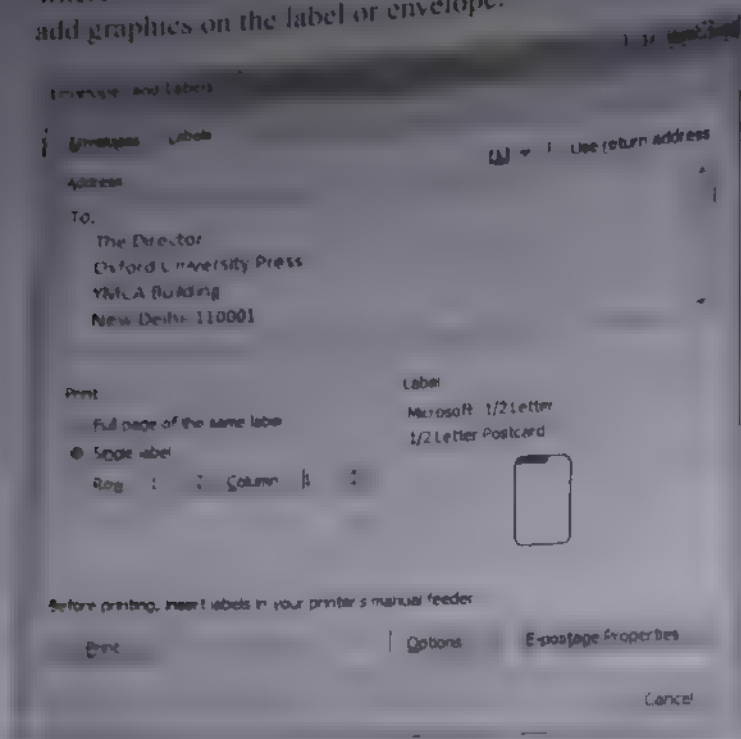


Figure 6.67 Labels

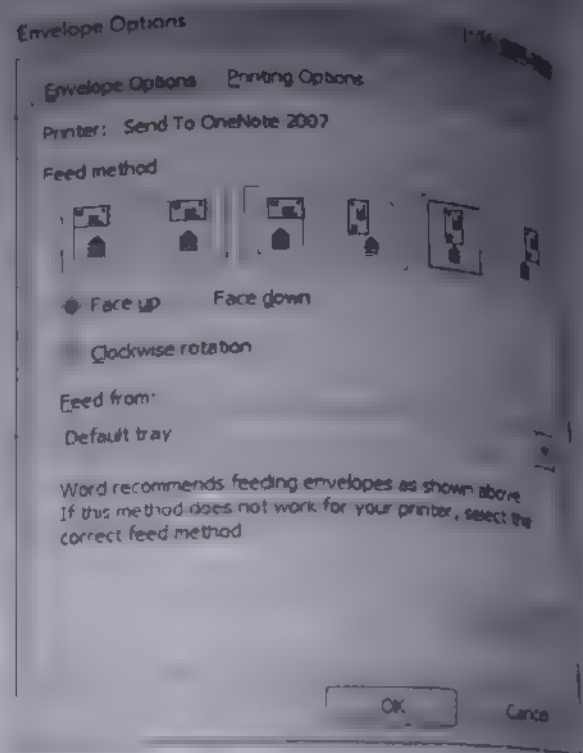


Figure 6.68 Envelopes

Summary

- Word 2013 is a word processing application software developed by Microsoft.
- Print Preview is used to see how a page would look when printed on a page.
- Editing a file means making changes in the file. While editing a document we may insert, delete, or replace some existing content.
- The simplest way to select text is to click on it and drag the mouse over the text you want to select.
- Alignment of text means how the text is displayed horizontally on the page. There are four options for text alignment: left, center, right, and justified.
- WordArt is a decorative text that you can add to a document.
- Microsoft Word provides bullets and numbers to maintain a list of items in a presentable order.
- Each cell can store some data.
- Microsoft Word allows you to apply beautiful borders on any or all of the four sides of a table. You can even add shading to its rows and columns to make the table more attractive.
- Headers and/or footers are added to include important information about the document.
- In Word, users can check the spelling and grammar of the text all at once by running the spelling and grammar checker.
- Word allows the AutoCorrect feature to correct typos, capitalization errors, and misspelled words. It is also used to automatically insert symbols and other pieces of text.
- Users can insert text automatically in Word with blocks of preformatted text from the AutoText gallery. Word also inserts text automatically when the user has typed only a few characters.
- Once an image is inserted, it is very difficult to move it exactly where you want. To move the picture freely, text wrapping feature is used.

Glossary

- Cell** The intersection of a row and a column.
Clipboard A small area to hold data temporarily.
Cursor The flashing vertical bar on the screen.
Document area Area where the text is typed.
Page margins Blank space around the edges of the page.

Resizing the table Changing the height and width of cells in a table.

Table A structure of vertical columns and horizontal rows.

Word processor Also known as word processing program, it is a software that processes words, paragraphs, pages, and entire papers.

Multiple-choice Questions

- Word is a _____ software.
 (a) database (c) spreadsheet
 (b) word processing (d) animation
- Select elements that cannot be inserted in MS Word.
 (a) Chart
 (b) Worksheet
 (c) PowerPoint presentation (d) None of these
- Select the most popular word processor.
 (a) Microsoft Word (c) AppleWorks
 (b) WordPerfect (d) OpenOffice.org
- _____ bar displays the name of the document.
 (a) Quick Access (c) Title
 (b) Ribbon (d) Scroll
- Title bar displays the name of the document and the _____.
 (a) program name
 (b) date of creation
 (c) date of last accessed
 (d) name of author
- _____ bar contains buttons for frequently used commands.
 (a) Quick Access (c) Title
 (b) Ribbon (d) Scroll
- The command to print a document is present in the _____ tab.
 (a) Home (c) Page Layout
 (b) File (d) View
- There are _____ rulers in a document.
 (a) 1 (c) 3
 (b) 2 (d) 4
- _____ is used to align text, graphics, tables, and other elements in a document.
 (a) Scroll bars (c) Ruler
 (b) File tab (d) Zoom Control
- Which view displays pages exactly as they will appear when printed?
 (a) Full Screen Reading
 (b) Web Layout
 (c) Outline
 (d) Page Layout
- Headers and footers cannot be seen in which view?
 (a) Draft (c) Outline
 (b) Web Layout (d) Page Layout
- _____ bar displays information about the document like the total number of pages and words.
 (a) Quick Access (c) Title
 (b) Ribbon (d) Status
- Name of a newly created document is _____ by default.
 (a) Word 1 (c) My Document
 (b) Document 1 (d) File 1
- In MS Word, a document is automatically saved with an extension _____.
 (a) .doc (c) .docx
 (b) .docs (d) .docz

15. Which key should be pressed to start a new paragraph?
 (a) Insert (c) Tab
 (b) Home (d) Enter
16. Print _____ is used to see how a page would look when printed on a page
 (a) Preview (c) Orientation
 (b) Layout (d) Window
17. If you want to add more than one space between two words, use the _____ key.
 (a) Home (c) Space Bar
 (b) Tab (d) Shift
18. Delete key is used to delete the character _____ the cursor.
 (a) in front of (c) above
 (b) behind (d) under
19. _____ button is used to reverse the last action(s) performed.
 (a) Cut (c) Undo
 (b) Paste (d) Redo
20. Arrow keys + _____ key is used to select text
 (a) F5 (c) F7
 (b) F6 (d) F8
21. To select a paragraph, click _____ times on the paragraph
 (a) 1 (c) 3
 (b) 2 (d) 4
22. To select the whole document, press Ctrl + _____ keys on the keyboard.
 (a) A (c) W
 (b) D (d) S
23. _____ command is used to move a text from one location to another
 (a) Cut Copy (c) Cut/Paste
 (b) Copy/Paste (d) Copy Copy
24. To paste text, press the _____ keys.
 (a) Ctrl + X (c) Ctrl + P
 (b) Ctrl + C (d) Ctrl + V
25. _____ button is used to enlarge the font size of the text
 (a) Large (c) Big
 (b) Grow (d) Increase
26. Each time the Shrink button is clicked, the _____ is reduced by _____ point(s).
 (a) 1 (c) 3
 (b) 2 (d) 4
27. The default font colour in MS Word 2013 is _____.
 (a) White (c) Black
 (b) Red (d) Grey
28. The default font in MS Word 2013 is _____.
 (a) Times New Roman
 (b) Comic Sans
 (c) Arial
 (d) Calibri
29. The default font size in MS Word 2013 is _____.
 (a) 10 (c) 12
 (b) 11 (d) 14
30. _____ determines the blank space on each side of a paragraph.
 (a) Alignment
 (b) Line spacing
 (c) Margin
 (d) Paragraph spacing
31. To get capital letter, press the letter along with _____ key.
 (a) Ctrl (c) Shift
 (b) Alt (d) Insert
32. _____ case option reverses the case of every individual character.
 (a) Upper (c) Sentence
 (b) Lower (d) Toggle
33. _____ stores data in a table.
 (a) Row (c) Cell
 (b) Column (d) Grid
34. While resizing a table, you can change the cell's _____.
 (a) height
 (b) width
 (c) both
 (d) none of these
35. To divide a cell into multiple cells, you will use the _____ command.
 (a) Merge (c) Split
 (b) Divide (d) Break

- has a list of typical misspellings and symbols
- AutoCorrect
- Spell Check
- Grammar Check
- Find and Replace
- checks for sentence fragments.
- AutoCorrect
- Spell Check
- Grammar Check
- Find and Replace
- can correct typos, capitalization errors, and misspelled words.
- AutoCorrect
- Spell Check
- Grammar Check
- Find and Replace
- feature inserts text automatically when the user has typed only a few characters
- AutoCorrect
- Spell Check
- AutoText
- Find and Replace
- is a series of commands that can be grouped together as a single command to accomplish a task automatically
- AutoCorrect
- AutoText
- AutoRun
- Macros
- To make the first character bigger than others, we use
- font size
- drop cap
- font style
- WordArt
- The options to add pictures, SmartArt, charts, and videos are present in the tab.
- Home
- Design
- Insert
- Page Layout
- feature is used to move a picture anywhere along with the text.
- AutoText
- Image Wrap
- Text Wrap
- Drop Cap

Answers to Multiple-choice Questions

- | | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1. (b) | 2. (d) | 3. (a) | 4. (c) | 5. (a) | 6. (d) | 7. (b) | 8. (b) | 9. (c) | 10. (d) |
| 11. (a) | 12. (d) | 13. (b) | 14. (c) | 15. (d) | 16. (a) | 17. (b) | 18. (a) | 19. (c) | 20. (d) |
| 21. (c) | 22. (a) | 23. (c) | 24. (d) | 25. (b) | 26. (a) | 27. (c) | 28. (d) | 29. (b) | 30. (c) |
| 31. (c) | 32. (d) | 33. (c) | 34. (c) | 35. (c) | 36. (a) | 37. (c) | 38. (b) | 39. (c) | 40. (d) |
| 41. (c) | 42. (b) | 43. (b) | | | | | | | |

Applications of Presentations

CHAPTER 7

Micro
display
founda

Syllabus Mapping	Unit
Basics of presentations slides, fonts, drawing, editing; Inserting: tables, images, texts, symbols, media, Design, Transition, Animation, Hyperlink, and Slideshow. Creating business presentations using above facilities.	Module II Unit 7

7.1 INTRODUCTION

PowerPoint is a presentation software developed by Microsoft for Microsoft Windows and Mac OS X. It supports text, shapes, graphics, pictures, and multimedia along with integration with other Microsoft Office products like Word and Excel. A document prepared in Microsoft Presentation is called presentation and is saved with the .pptx extension, whereas the file extension of the prior versions of PowerPoint was .ppt.

Microsoft (MS) PowerPoint is called a presentation software because it helps users to prepare beautiful presentations for different purposes. For example, teachers make presentations to explain concepts and teach students in a class, managers make presentations to present their ideas, introduce a new product, or to explain an organizational structure, etc. In this chapter, we will learn how to draw beautiful presentations using MS PowerPoint.

The chapter starts with an introduction of the MS PowerPoint software and deals with the basics of creating, opening, and saving a PowerPoint presentation. We will learn how to apply attractive backgrounds and templates, and insert text, Word tables, Excel sheets, pictures, clip art, word art, movies, sound, headers, footers, and other objects. We will learn how to insert or delete slides from an existing presentation. The chapter also discusses how to view a presentation, printing slides, and adding animations and transition effects.

7.2 BASICS

Before making attractive presentations, let us first learn how to open the presentation software, how to open an existing presentation, and save a new or an existing presentation.

7.2.1 Using PowerPoint

To use Microsoft PowerPoint, you must first open it. Follow the steps given below to open this software.

Step 1: Click on Start button.

Step 2: Click on All Programs.

Step 3: Click on Microsoft Office.

Step 4: Click on Microsoft PowerPoint 2013.

Microsoft PowerPoint 2013 is then started and the presentation window (as shown in Fig 7.1) is opened. Now let us understand the different components of the presentation window that forms the foundation of the MS PowerPoint 2013 software.

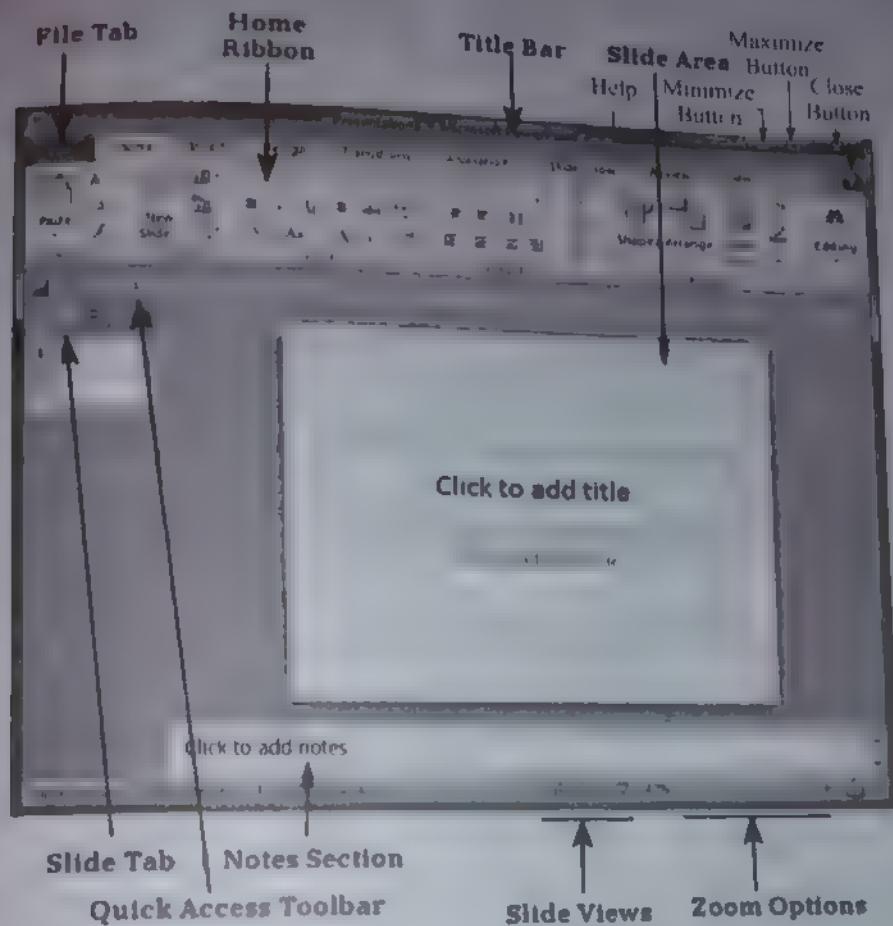


Figure 7.1 PowerPoint window

Components of Presentation Window

File Tab The File tab allows you to save presentations, open existing ones, create new presentations, and print them. Other file related operations can also be executed from this view.

Ribbon The ribbon (Fig. 7.2) has the following components:

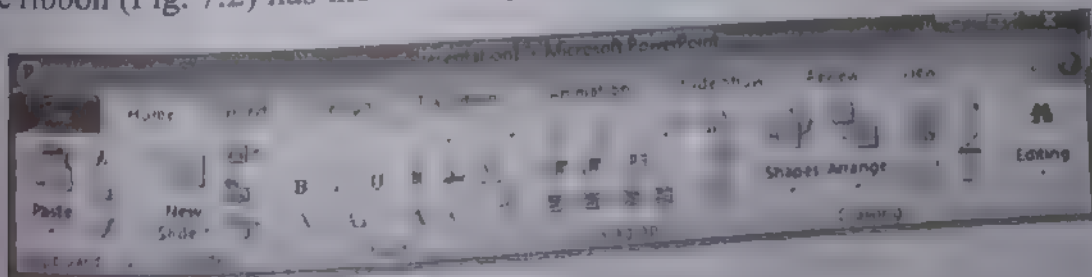


Figure 7.2 Ribbon

Tabs Tabs appear across the top of the Ribbon and contain groups of related commands. Some tabs in PowerPoint are Home, Insert, Design, Animations, etc.

Groups Every tab has a number of groups. These groups contain related commands. The name of the group appears below the group on the ribbon. For example, in the Font group all commands related to font type, size, and colour are present.

Title Bar The Title bar is present at the top of the window. It displays the name of the document and name of the program (which is Microsoft PowerPoint).

Slide Area It is the area where the slide is created and edited. You can insert, edit, and delete text, images, shapes, and multimedia in this area.

Help The Help icon (?) opens the in-built help of PowerPoint and is used to get help related to topics in PowerPoint. The Help window (Fig. 7.3) displays a list of common topics which you can browse to find details about a topic of your choice. You can even search for specific topics from the Search bar at the top.

Zoom Options Zoom control allows you to zoom in for a closer look at the text. It consists of a slider that can be slid left or right to zoom in or out. You can click on the – and + buttons to increase or decrease the zoom factor. The maximum zoom supported by PowerPoint is 400%. 100% zoom is indicated by the mark in the middle.

Slide Views The Slide Views are a group of four buttons present to the left of the Zoom control, near the bottom of the screen. It helps you to switch between PowerPoint views. The different views are

Normal Layout View As the name suggests, this view displays the page in normal view with the slide on the right and a list of thumbnails to the left. It is useful while editing individual slides and rearranging them.

Slide Sorter View In this view, all the slides are displayed as a matrix. It is useful when you just have to rearrange and not edit the slides.

Reading View This view is similar to a slide show. Like in case of a slide show, you cannot edit the slide in this view and by clicking on a slide you will move the presentation to the next slide. However unlike a slide show which always starts from the first slide, in the Reading View you can view any particular slide.

Notes Section This section allows you to add notes to the presentation. These notes will not be displayed on the screen during the presentation but serve as quick reference for the presenter.

Quick Access Toolbar It is present above the File tab and to the left of the Title bar. The Quick Access Toolbar contains buttons for frequently used commands (like Undo, Redo, Save, Print Preview). You can customize this toolbar to contain commands that you use frequently.

Slide Tab The Slide tab can be seen only in the Normal view. It displays all the thumbnails of all the slides in sequence. You can add, delete, and rearrange slides from this section.

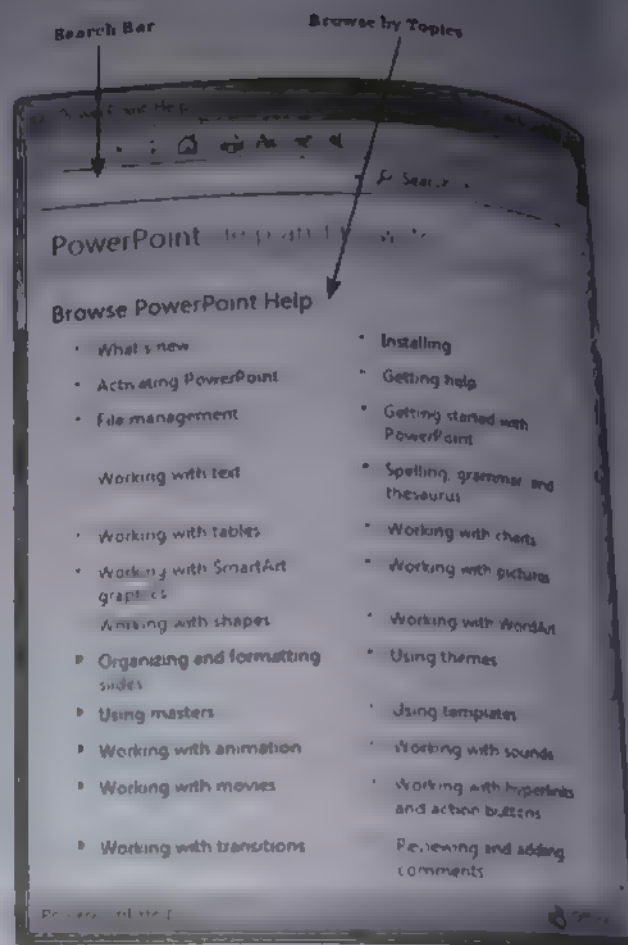


Figure 7.3 Help window

Minimize Button It is represented as an underscore in the top-right corner of the window. When a window is minimized, it will remain on the taskbar, but not visible.

Maximize Button It is used to enlarge a window, to make it fill the entire screen or the program window in which it is contained.



A maximized window cannot be moved until its size is reduced.

Close Button It is used to close a window. When the user clicks on this button, only the current window is closed. That is, in case he had opened many windows of a particular application, closing one window does not exit the application.

7.2.2 Opening a PowerPoint Presentation

To open a PowerPoint presentation, follow the steps given below.

Step 1: Click on the File tab.

Step 2: Click on Open option.

Step 3: From the Open dialog box that appears on the screen, browse the presentation you want to open.

Step 4: Click on Open as shown in Fig. 7.4

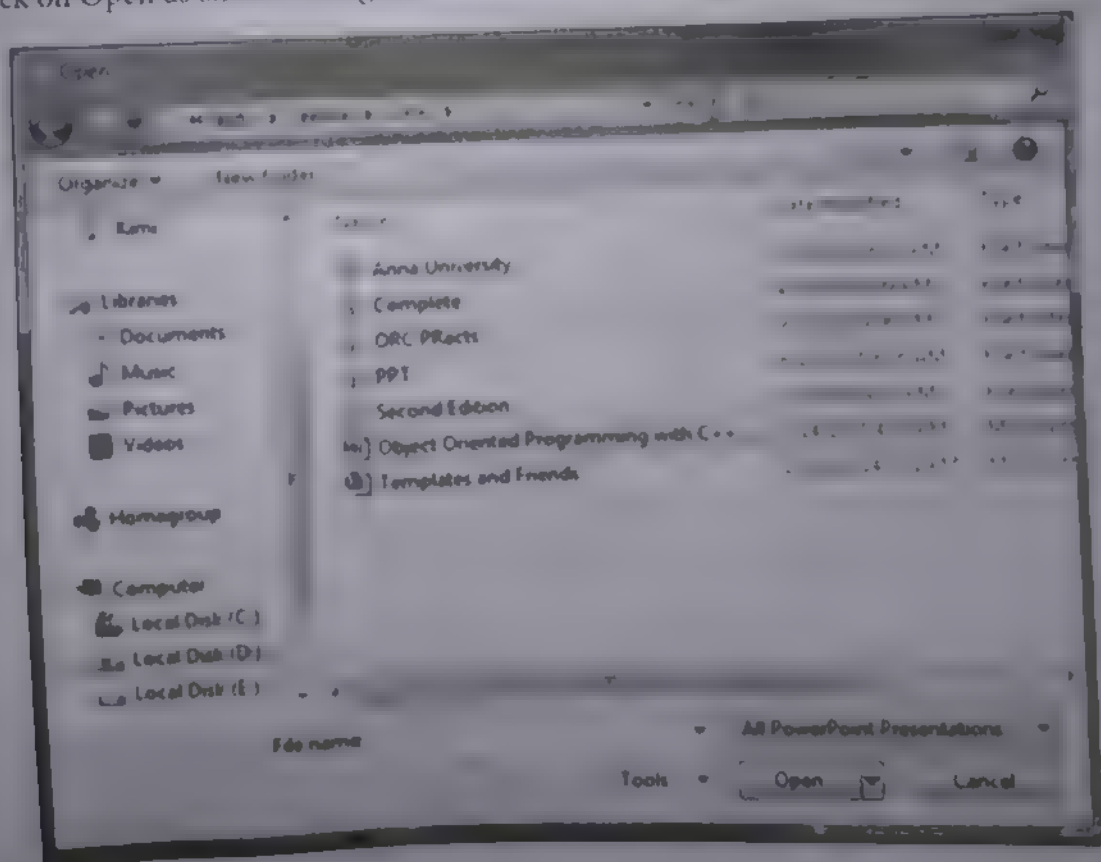


Figure 7.4 Open dialog box

7.2.3 Saving a Presentation

To save a PowerPoint presentation, follow the steps given below:

Step 1: Click on the File tab.

Step 2: Click on Save as option.

Step 3: From the Save as dialog box that appears on the screen, browse the location where you want to save your presentation.

Step 4: Type a suitable name for your presentation.

Step 5: Click on Save as shown in Fig. 7.5.

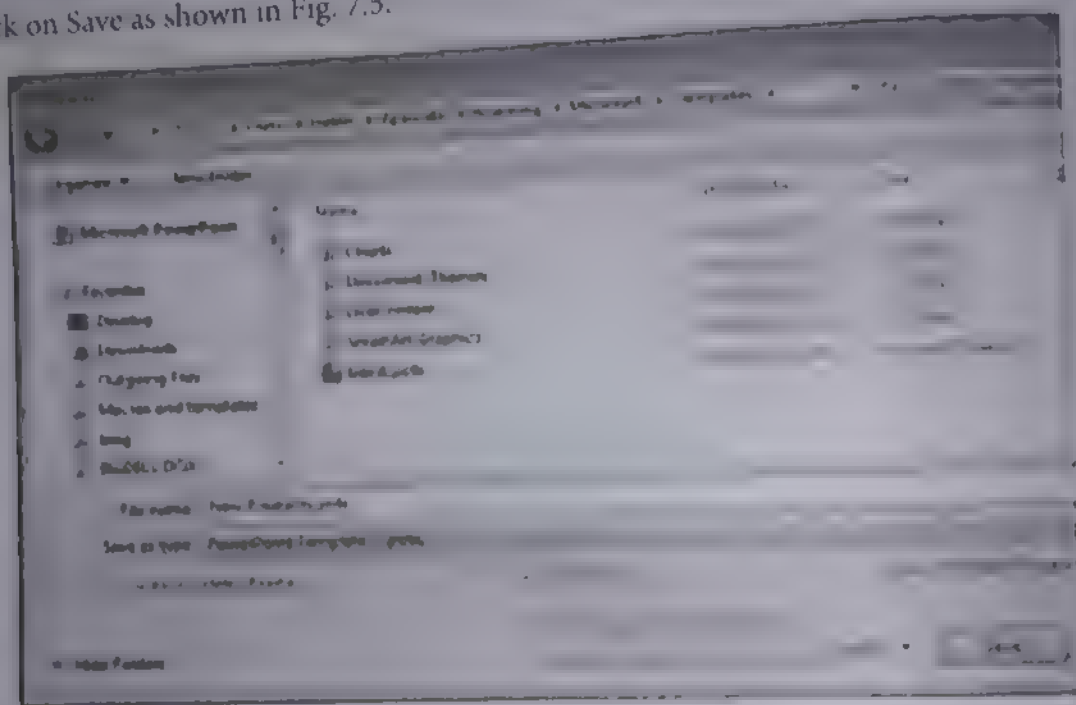


Figure 7.5 Save dialog box

Once the presentation is saved, you will have to save it every time you edit it. For this you must click on the Save icon in the Quick Access Toolbar or click on the Save option from the File tab.

7.3 CREATION OF PRESENTATION

PowerPoint has many tools to help you create a beautiful presentation. These tools are organized logically into various ribbons. In this section we will learn how to create a new presentation and add or edit text in it.

7.3.1 Creating a Presentation Using a Template

To make your presentation contain thoughtful arrangement of elements and colour, fonts, effects, style, and layout, you can create the presentation using a template.

PowerPoint 2013 allows you to apply already installed templates, apply your own custom templates, or to search from a variety of templates available on Office.com. To apply a template, follow the steps given below:

Step 1: Click on File tab and select the New option.

Step 2: Under Available Templates and Themes (Fig. 7.6), you can do any of the three things.

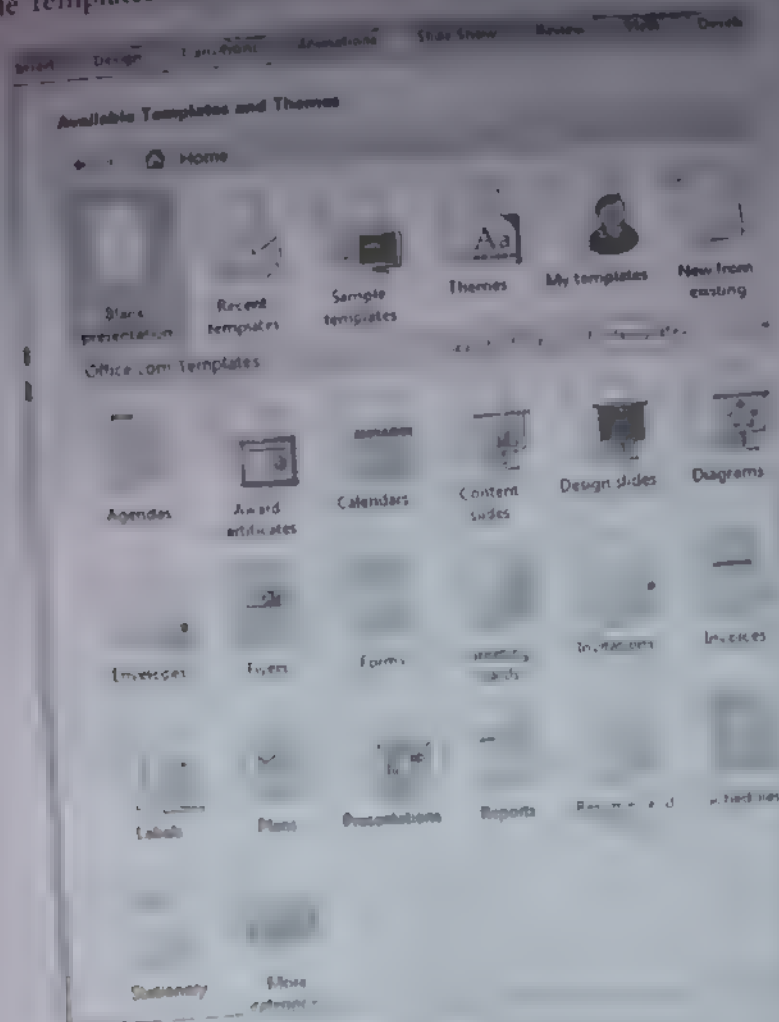


Figure 7.6 Available templates and themes

- First, click on Recent Templates to reuse a template that you had recently used. Click on the desired template and then click on Create
- Second, to use an already installed template click on My Templates, click on the desired template and then click on OK.
- Third, to use an absolutely new template, you can search on Office.com. To do so, click a template category. Select the template that you want, and then click on Download to save the template from Office.com to your computer

7.3.2 Creating a Blank Presentation

Follow the steps given below to create a blank presentation:

Step 1: Click on the File tab.

Step 2: Click on New.

Step 3: Select Blank presentation. It will be highlighted by default.

Step 4: Click on Create. A new blank presentation will be displayed on the screen as shown in Fig. 7.7.

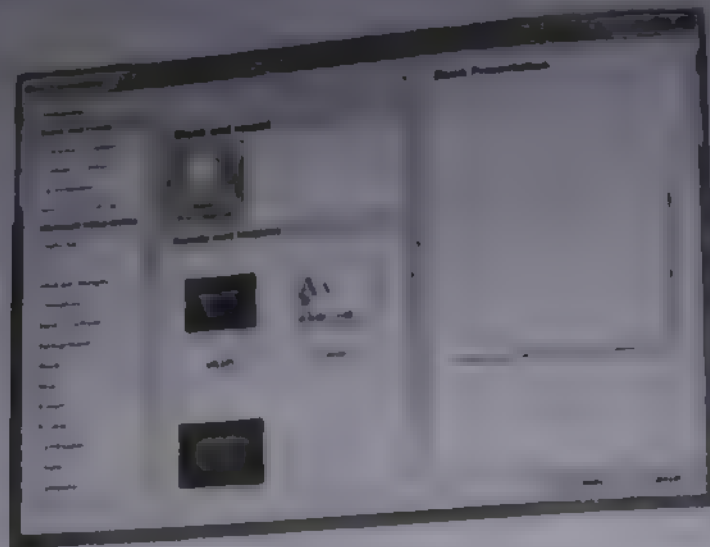


Figure 7.7 Creating blank presentation

7.3.3 Entering and Editing Text

To enter text in a slide, simply type the text in the placeholder. To enter text outside a placeholder, you must insert a text box and then type in it. A text box can be added by following the steps given below.

Step 1: Click on the Insert tab.

Step 2: From the Text group, click on Text Box and while holding down the left mouse button drag the mouse to draw the text box.

Step 3: Type the text in it as shown in Fig. 7.8.

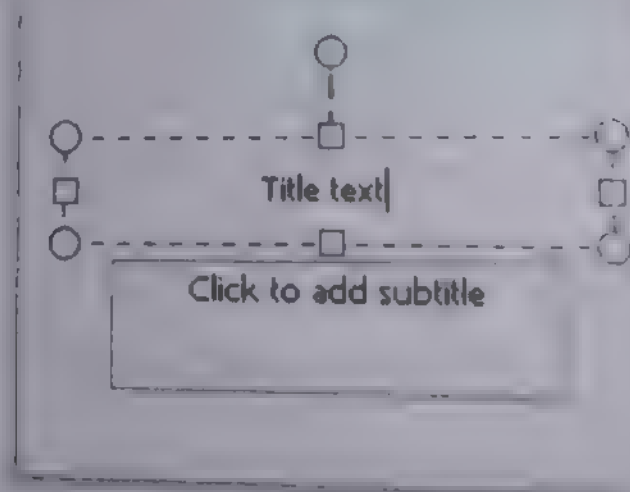


Figure 7.8 Entering text

Step 4 (optional): To format the text box, right click on it and select Format Shape. Now you can choose fill color, style its border, choose a line color, align text, add 3D effects and a lot more. To edit the text, just click in the placeholder or the text and type the text.

7.3.4 Inserting and Deleting Slides in a Presentation

Many a times we may find the need to add a new slide in an existing presentation. To add a slide, follow the steps given below:

Step 1: Right click in the Navigation Pane under the slide where you want to insert the new slide

Click on the New Slide option as shown in Fig. 7.9

A new slide is inserted. The layout of the newly added slide can be changed to suit your design

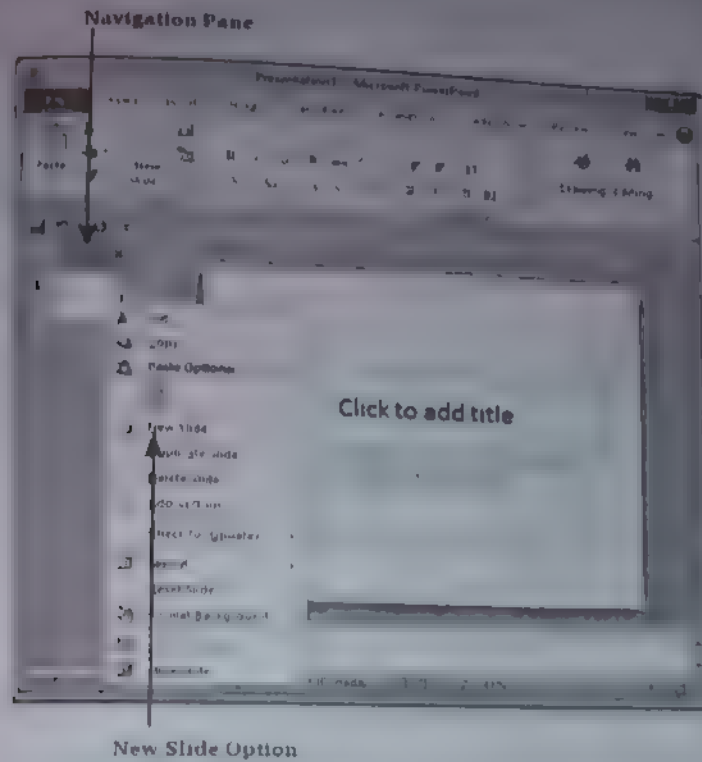


Figure 7.9 Inserting new slide

To change the slide layout, right click on the slide and select Layout and choose from the existing layout styles as shown in Fig. 7.10.

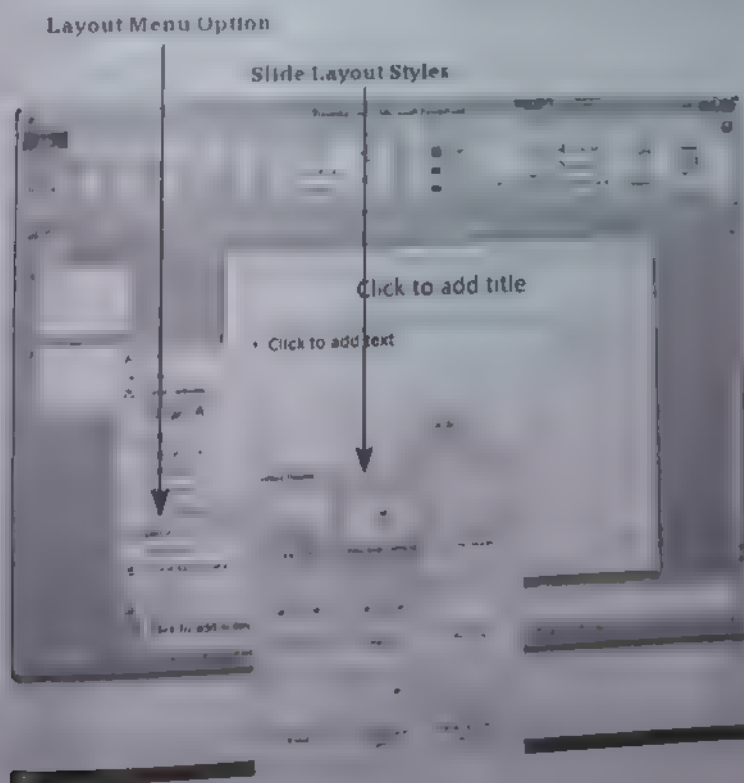


Figure 7.10 Slide layout

While working on presentations, at times you may feel the need to delete one or more slides. You can delete slides from both Normal view and Slide Sorter view. To delete a slide, simply select the slide and press the Delete key. You can also right click on the slide and then press the Delete Slide option as shown in Fig. 7.11.

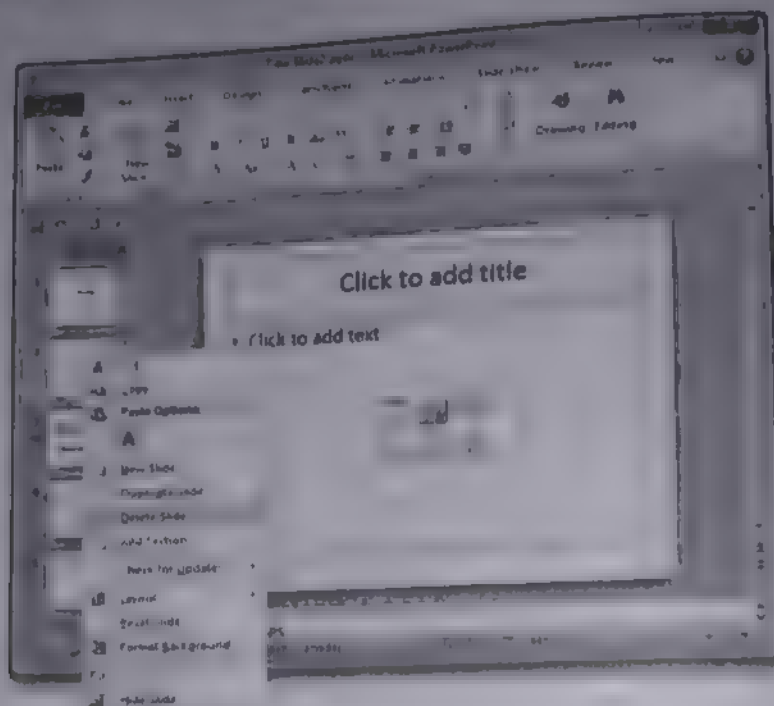


Figure 7.11 Deleting a slide

7.4 PREPARATION OF SLIDES

In this section, we will learn about adding some objects like pictures, tables, worksheets, etc. in the slides.

7.4.1 Inserting Word Table or an Excel Worksheet

To insert an Excel worksheet in your PowerPoint slide, follow the steps given below:

Step 1: Click on the Insert tab.

Step 2: From the Text group, click on Object.

Step 3: From the Insert Object window, select Create from File as shown in Fig. 7.12.

Step 4: Click on Browse, and in the Browse dialog box, select the Excel workbook.

Step 5: Click on OK.

Once the worksheet is inserted, you can expand its view by resizing and placing it anywhere in the slide.



To insert an existing worksheet, select Create from File option in Step 3. Browse the file and then click on OK.

An alternate way to insert an Excel sheet is to go to the Insert tab. Click on the down arrow under Table. From the menu that pops up select Excel Spreadsheet.

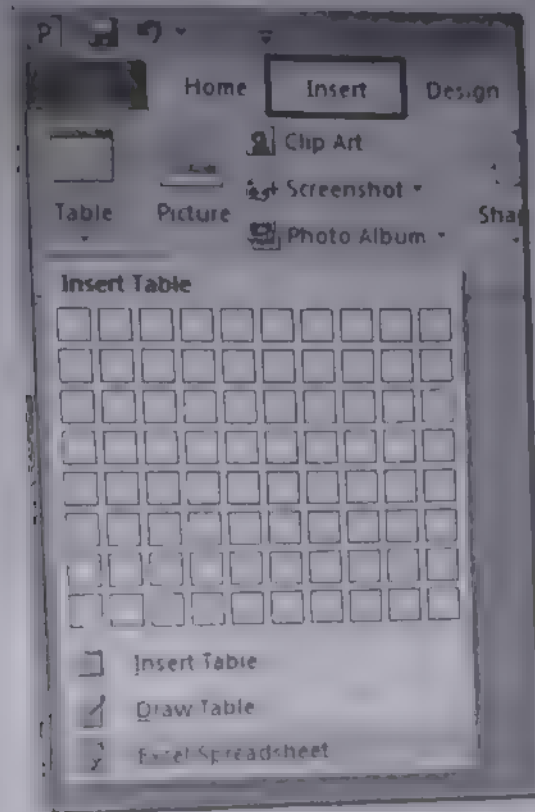


Figure 7.12 Inserting an Excel worksheet

To insert a table from Word into a PowerPoint slide, follow the steps given below

- Step 1:** Select the Table in the Word document
- Step 2:** From the clipboard group, click on Copy.
- Step 3:** Open the Slide. From the Clipboard group, select Paste.

7.4.2 Inserting Pictures in Slides

It is said that a picture is worth a thousand words. We understand more clearly by looking at pictures than reading the text. Pictures also make the text interesting. Therefore, PowerPoint allows you to insert pictures in the slides in two ways. First, by inserting a picture from a file stored on your computer and second, by inserting a picture from the web. In this section, we will read about both the ways.

Follow the steps given below to insert a picture from a file on your computer

- Step 1:** Select the Insert tab.
- Step 2:** Click on Pictures. A dialog box appears on screen as shown in Fig. 7.13 (a).
- Step 3:** Select the picture.
- Step 4:** Click on Insert.
- Step 5:** Drag the picture to move it at the desired position.

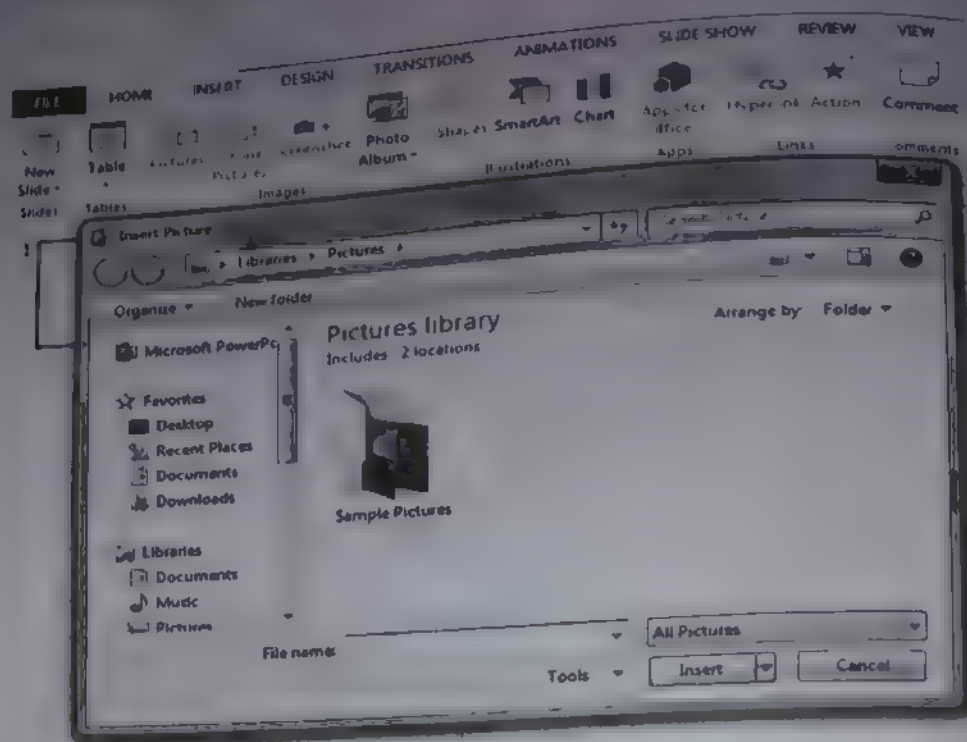


Figure 7.13 (a) Inserting picture from computer

Insert Online Pictures

If a picture is not there in your computer, then you can take it from the Internet using the Insert Online Picture

Step 1: Select the Insert tab.

Step 2: Click on Online Pictures.

Step 3: Enter the name of the picture you want to search for.

Step 4: Select the image you wish to insert and click on Insert (Refer Fig. 7.13 (b)).

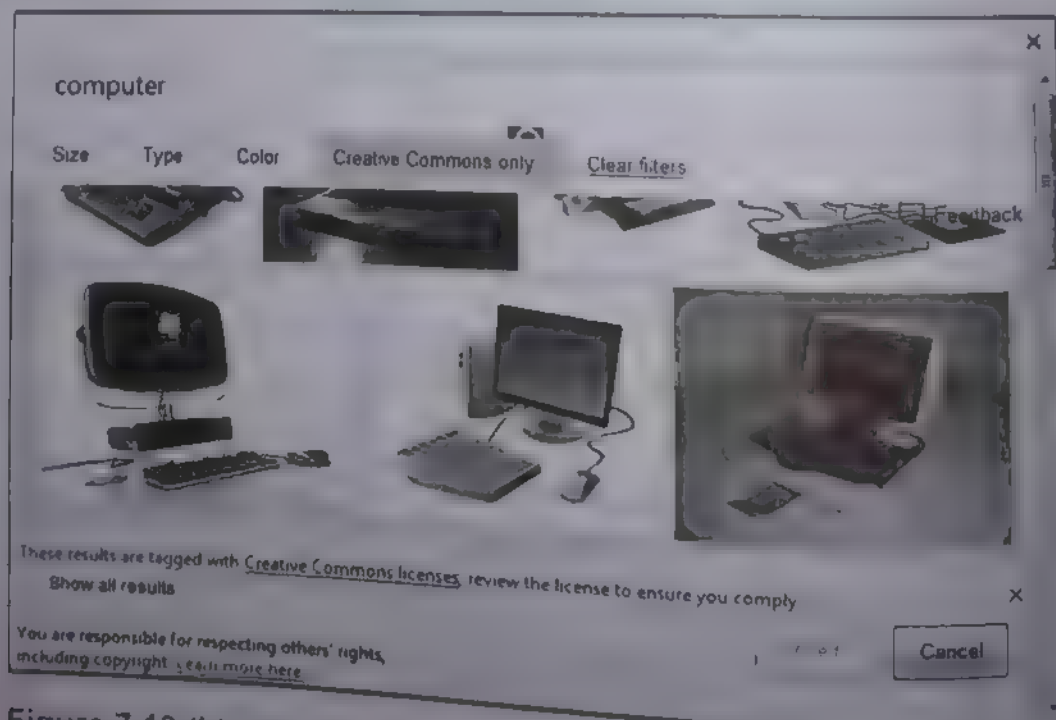


Figure 7.13 (b) Inserting picture from Internet

7.4.3 Inserting Other Objects

PowerPoint allows users to create animated presentations containing text, images, sounds, and other media objects. Besides these custom features, you can also add linked objects in your slides. In this section, we will see how to insert linked objects in your presentation. While objects simply embedded in your presentation remain unchanged when the source file changes, linked objects reflect any changes made to the source file. That is, if you modify the data in a worksheet (using Excel) that is linked to your presentation, the data will be automatically updated in the PowerPoint. Follow the steps given below to insert a linked object in your PowerPoint presentation.

Step 1: Click on the Insert tab.

Step 2: Under the Text group, click on Object.

Step 3: Select Create new or Create from File (Refer Fig 7.14).

Step 4: Browse the file from your computer. Click on Open (Refer Fig 7.15).

Step 5: Check the box next to the word Link. This will insert the file as a linked object.

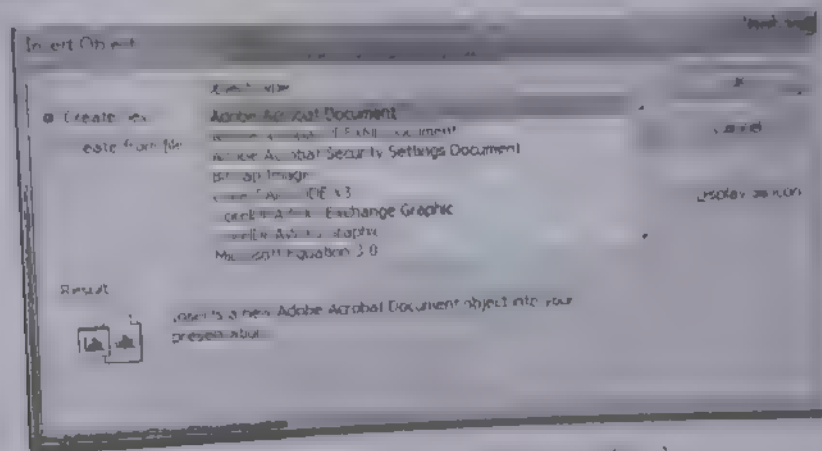


Figure 7.14 Insert object (Create new option)

Step 6: Check the box next to Display as icon. This is optional. Do this if you want the linked object to appear as an icon in your PowerPoint presentation until you click it. If you do not check this box then the content of the file would appear in your presentation.

Step 7: Click on OK.

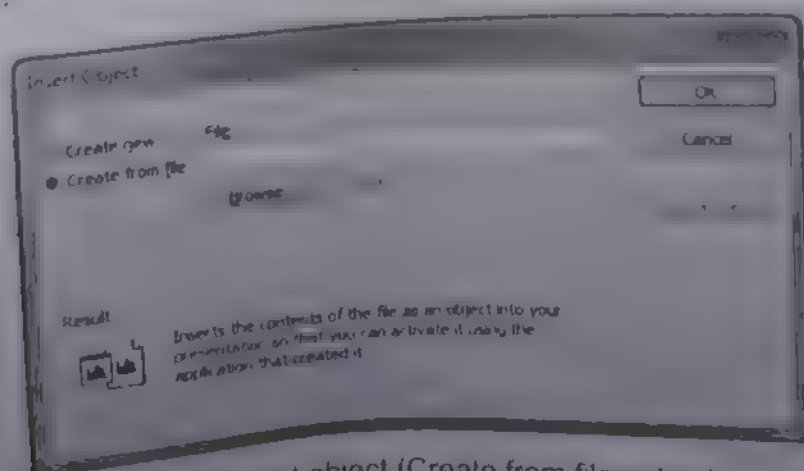


Figure 7.15 Insert object (Create from file option)

An alternative way to insert chart, table, smart art graphics, picture from file, clip art or any media file like audio or video is to click on the desired option available in the placeholder as shown in Fig. 7.16

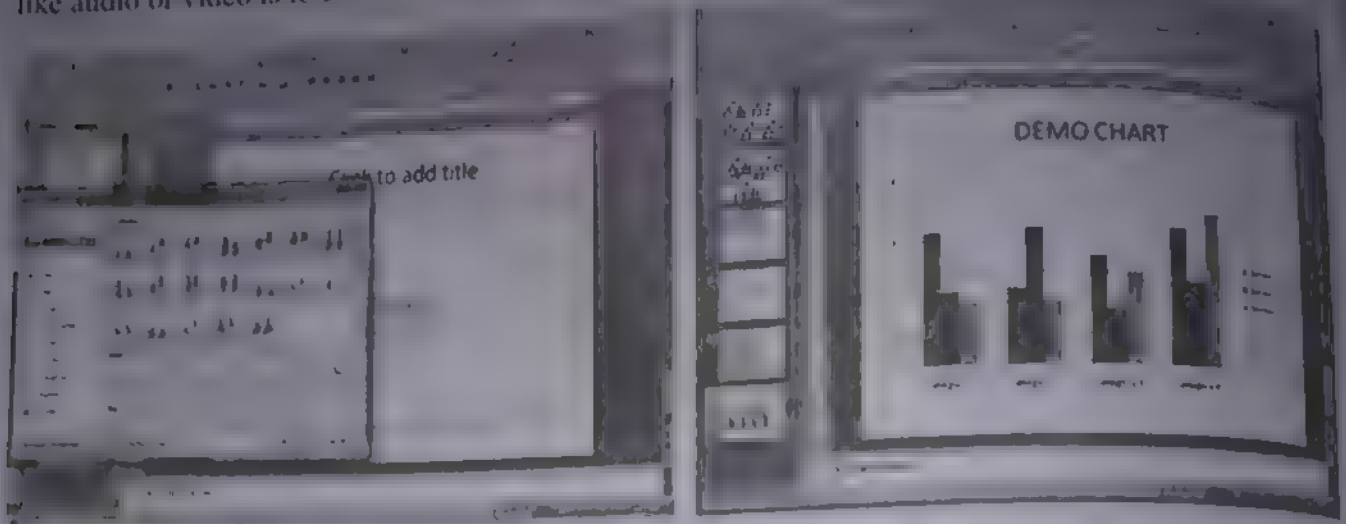


Figure 7.16 Inserting charts in a slide

7.4.4 Resizing and Scaling an Object

Resizing and scaling stretches or shrinks the dimensions of an object. To resize or scale a picture, follow the steps given below:

Step 1: Click on the object that you want to resize.

Step 2: To increase or decrease the size in one or more directions, drag the sizing handle away from or toward the center (Refer Fig. 7.17).

To resize an object to an exact height and width, follow the steps given below

Step 1: Click on the object that you want to resize.

Step 2: Click on the Format tab.

Step 3: In the Size group, enter the desired measurements in the Height and Width boxes (Refer Fig. 7.18)



Figure 7.17 Resizing an object with a sizing handle

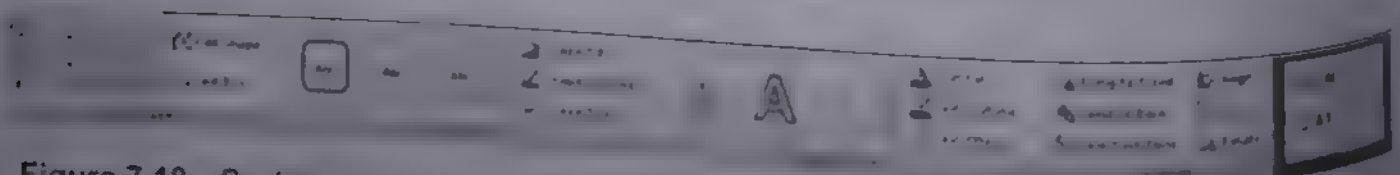


Figure 7.18 Scaling an object

7.5 PROVIDING AESTHETICS

In this section, we will read about some interesting features in PowerPoint. We will be adding beautiful effects to our plain text, add some styles and colours to simple lines, and insert movies, sounds, headers, and footers.

7.5.1 Enhancing Text Presentation

Writing plain text becomes very boring to read. To present your ideas, thoughts, and content effectively you must make your text attractive. For this, you must format your text by selecting different font sizes, color, and style. Follow the steps given below to format the text.

Step 1: Select the text.

Step 2: Click on the appropriate button to make the desired changes (Refer Figs 7.19 and 7.20).

Step 3 (Optional): To apply more effects, click on the Format tab, then from the Word Art styles group, click on any text effects (Refer Figs 7.21 and 7.22).

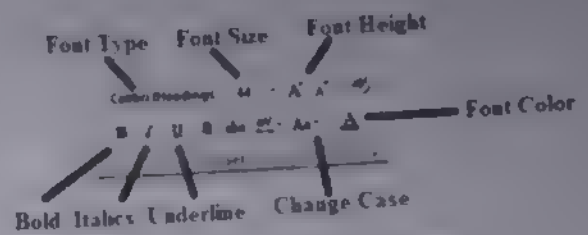


Figure 7.19 Working with fonts

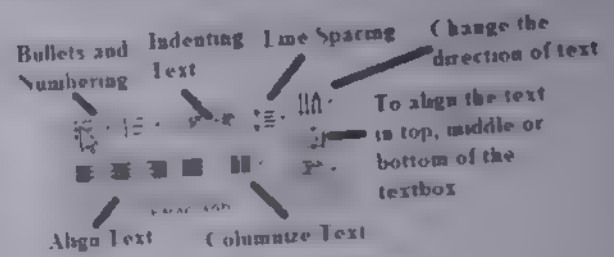


Figure 7.20 Formatting text

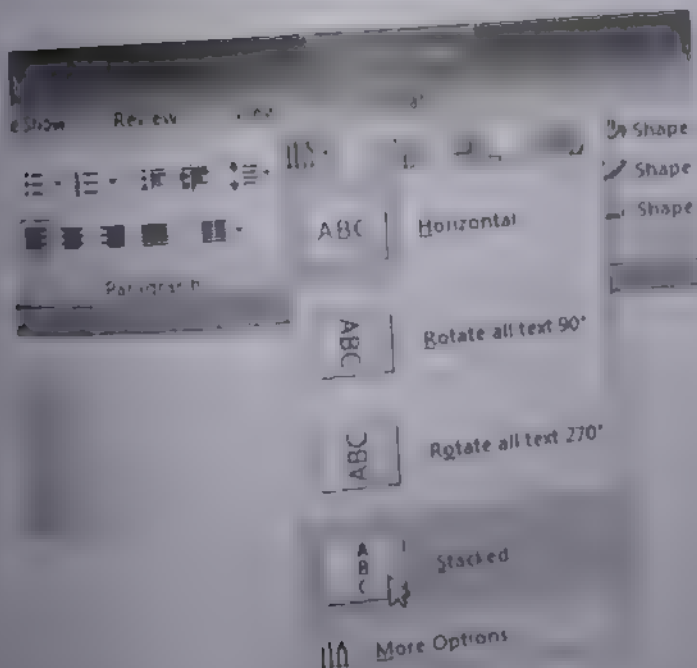


Figure 7.21 Applying text effects



Figure 7.22 Working with Word Art

7.5.2 Working with Color and Line Style

You can apply different styles on a line to include colors, shadows, gradients, and three-dimensional (3-D) perspectives. To apply different shape effects to a line, follow the steps given below:

- Step 1:** Select the line that you want to change. To change multiple lines, select the first line, and then press and hold Ctrl while selecting the other lines.
- Step 2:** Click on the Format tab (Refer Fig. 7.23).
- Step 3:** In the Shape Styles group, click on the small arrow next to Shape Effects and select the desired effect.

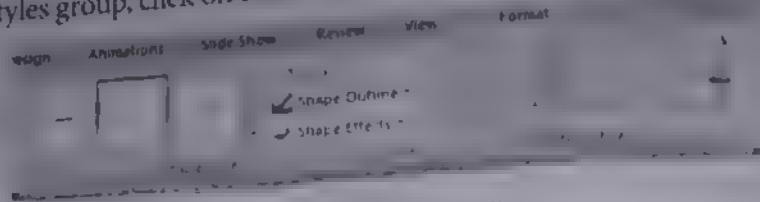


Figure 7.23 Applying effects to a line

To change the color and style of the line, follow the steps given below:

- Step 1:** Select the line.
- Step 2:** Click on the Format tab.
- Step 3:** Click on the small arrow next to the Shape Outline (Refer Fig. 7.24).
- Step 4:** Click on the desired color. The color of the line will be changed.

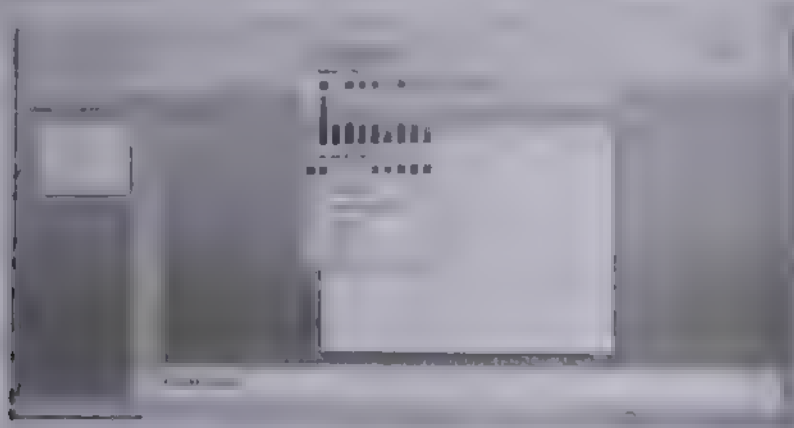


Figure 7.24 Changing the color of the line

- Step 5:** To change the style of the line, click on the small arrow next to Shape Outline in the Shape Styles group.
- Step 6:** Select the desired style as shown in Fig. 7.25. The desired style will be applied.

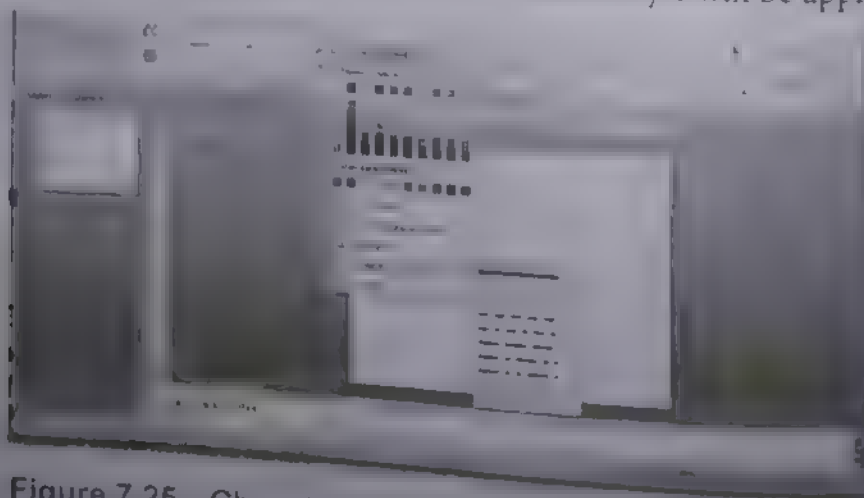


Figure 7.25 Changing the style of the line

7.5.3 Adding Movie and Sound

PowerPoint 2013, you can add movies (also known as video files) with formats such as AVI (.avi), Windows Media Video (.wmv) and MP4 (.mp4 and .mpe). Whether you insert your audio/sound (.mp3, .wma, .wav) from a file in your computer or from Clip Organizer, it will be run during the slide show.

You can insert an audio or video that is relevant to your topic and audience. This would help you to attract audience's attention and also eliminate the need for you to jump between your PowerPoint presentation and media files.

To insert a movie or sound, follow the steps given below:

- Step 1:** In Normal view, Click on the Insert tab.
- Step 2:** Click on the small arrow next to Video or Audio (whichever you want to insert) (Refer Fig. 7.26).
- Step 3:** If you want to add from a file on your computer, click on Movie/Sound from File option.
- Step 4:** The Insert Sound/Movie dialog box will open. Browse the computer to locate the file.
- Step 5:** Click on Open.
- Step 6:** In case you insert a sound file, the Sound icon will appear on the slide. Drag it to any location on the slide.
In case you insert a video file, click on the video clip area and use the resize handles to resize it, if necessary.
- Step 7:** You will see an Options task. Use this tab to get a preview or control the volume and other features.



The audio and video files will be played during the slide show.

Figure 7.26 Adding a video

7.5.4 Adding Headers or Footers

When headers or footers are added to include important information about the slides. This information includes the name of the presentation, the date on which the presentation was created or delivered, name of the organization, etc. Headers and footers are mainly used for two reasons in a PowerPoint presentation.

- First, they are used to display recurring information (like slide number) on the slides to help the audience keep track of the presentation content.
- Second, they help in creating a branding effect by keeping your name or the company's name in front of the audience.

Headers and footers appear on every slide. Before using them, make sure they don't make your presentation look cluttered. The default font size is usually small enough to minimize distraction, but you can always change their font size and placement to minimize cluttering.

There is no general rule as to what should come in header and what should come in footer. Follow the steps given below to insert headers and footers in your presentation:

- Step 1:** Click on the Insert tab.
- Step 2:** In the Text group, click on Header and Footer.
- Step 3:** In the Header and Footer dialog box, go to the Slide tab.
- Step 4:** To make Footer, select the Footer check box (Refer Fig 7.27).
- Step 5:** Type the name of the company or any other text that you want to appear in the Header/Footer or select the information you want to include.

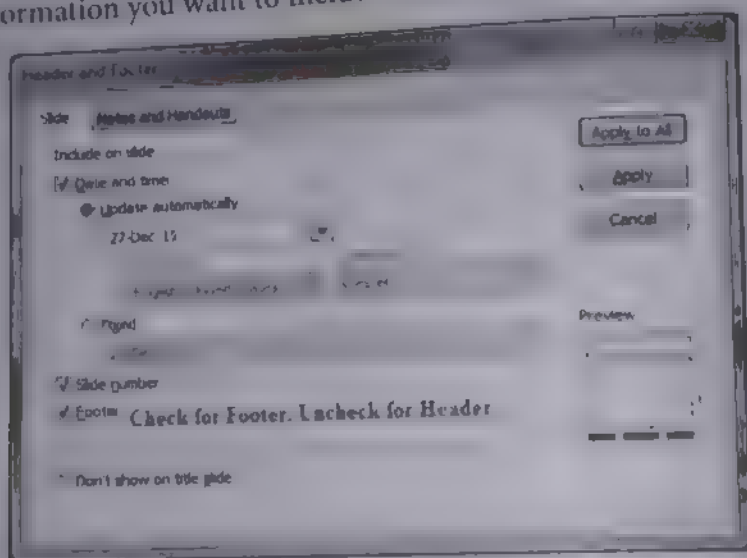


Figure 7.27 Adding Header/Footer

Select the Date and Time check-box to make the date appear on the slide(s). You can choose whether you want the date and time to Update Automatically, or set them to Fixed. You will get different presentation options in the former case but in Fixed option, you will simply get a text box where you will have to type the date.

- Step 6:** To exclude, header and footer from the Title slide, check the Don't show on title slide box.
- Step 7:** To display header/footer information on only the selected slide, click on Apply and to display on all of the slides in your presentation, click on Apply to All.

7.6 PRESENTATION OF SLIDES

Once all the slides are ready, you can go further preparing for presentation. In this section, we will discuss viewing presentation, setting up presentation, and printing slides and handouts.

7.6.1 Viewing a Presentation

PowerPoint has several views to work on presentation. We have already seen some views in this chapter. In this section, we will see how to view a presentation in the Notes Page view (refer Fig. 7.28). This view displays one slide at a time with an area showing the slide's notes at the bottom and is an excellent

view to use if you have a lot of notes to type in or view (refer Fig 7.29). You can zoom the slide out if it is hard to read those notes.

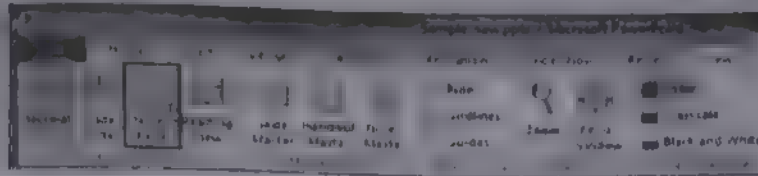


Figure 7.28 Notes page command in View tab

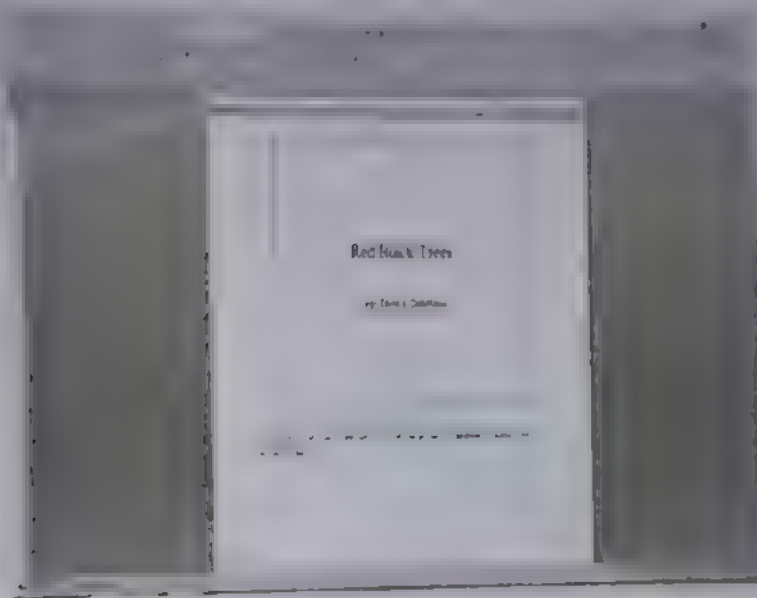


Figure 7.29 Notes displayed using the Notes Page View

The Notes Page view also helps users to view your own notes in a page view. Each slide created contains its own note page. These pages can be printed to be used as a reference while delivering the presentation. The notes do not show on the screen during Slide Show view.

To access the Notes Page view, click on the View tab. From the Presentation Views group click on the Notes Page button. The upper box of the page will display a single slide and the lower box of the page is the Notes area.

The Zoom slider is placed towards the right on the Status bar to get a better, zoomed-in view of any part of the Notes. Putting notes in the Notes section of either Notes Page view or Normal view is an excellent way of including reminders about topics that you have to discuss while delivering the presentation.

7.6.2 Choosing a Set-up for Presentation

PowerPoint has various options for setting up and playing a slide show. For example, you can set up a presentation that can be displayed at a kiosk and make your slide show repeat with continuous looping. For this, we should use the Set Up Show feature which has various settings that determine what happens when you go into Slide Show view and deliver your presentation.

Follow the steps given below to use the Set Up Show feature.

Step 1: Click on Slide Show tab

Step 2: From the Set Up group, click on Set Up Slide Show. The dialog box with many options will be opened (Refer Fig. 7.30).

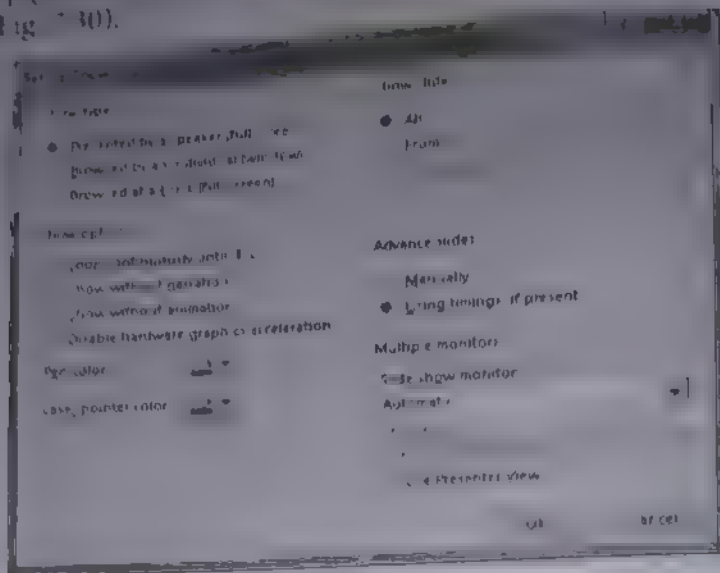


Figure 7.30 Choosing a set up

Let us understand some important aspects of this window.

Presented by a Speaker (Full Screen) It is the default option in which the slide show view is full screen and you can click from slide to slide.

Browsed by an Individual (Window) In this, Slide Show view is not a full screen but a resizable window. You can click from slide to slide.

Browsed at a Kiosk (Full Screen) In this, slide show is full screen but you can NOT click from slide to slide. For such a show automatic timing is set to move from one slide to another. You can even create action buttons or other hyperlinked objects to allow navigation from one slide to another. This type of presentation is called a self-running presentation.

Loop Continuously until 'Esc' It continuously loops your presentation and is extensively used for a trade show.

Show without Narration It is used if you are using PowerPoint's narration feature. Usually in self-running presentation, you may have narrated the presentation but if you want that in your presence, the narration must not be used. In such a situation, this option is very useful.

Show without Animation This option is used when you have a presentation that may have animation that you want to use only with certain audiences or in certain situations. In such a situation, you can turn off all animation.

Disable Hardware Graphics Acceleration Hardware acceleration helps animation and video run more smoothly. Sometimes you often get a message to upgrade your graphics card or its drivers. It would always be better if you do so but in case you are not able to do that, then to stop such messages check this box.

Pen Color This option is used to set the pen color for drawing during Slide Show view.

Laser Pointer Color It is used to set the laser pointer color.

Choose Which Slides You Want to Show It is used to specify which slides you want to show. The default option is to show all slides. An alternate way to do so is hide the slide(s) by right clicking them in the left-hand pane and clicking on the Hide Slide option.

Use Auto
default au
disable th
Configu
projector
Show vi
The L

7.6.3

You can
or custo
depend
a piece

Step 1:

Step 2:

Step 3:

Step 4:

Step 5:

Step

Step

Fig

Fig

- **Automatic Timing or Not** Usually, in self running presentation, automatic timing is set. By selecting automatic timing will be used if it exists but if you are doing a live presentation then you can select this timing by choosing the Manually option.
- **Configure Multiple Monitors** This option is used when you want to attach a second monitor or an LCD projector to your computer. From the Slide Show monitor drop-down list, you can choose where Slide Show view appears. From the Resolution drop down you can select a resolution.
- The Use Presenter View box is checked to see a special Presenter View on the primary monitor.

7.6.3 Printing Slides and Handouts

You can take printouts of slides in PowerPoint 2013. You can either take handout of the entire slide show or customize the number of slides that are to be printed out on each page. The number of slides on a page depends upon how detailed each slide is. Our main aim should be to print as many slides as possible on a piece of paper provided that everything on the slide is at a legible size.

Step 1: Open the slide show for which you want to print a handout.

Step 2: Click on the File tab.

Step 3: Click on the Print option as shown in Fig. 7.31.

Step 4: Click the Full Page slides from Print Layout.

Step 5: Select the number of slides as shown in Fig. 7.32 in the Handout section to select the number of slides you want on each page.

Click on the High Quality option box to ensure that images in your presentation look better. For this use more ink to print. This option is used when your slide show includes a number of highly colorful and detailed slides with important images or graphs.

Step 6: Select the number of copies of your handout that you want to print from Copies.

Step 7: Click on the Print button.

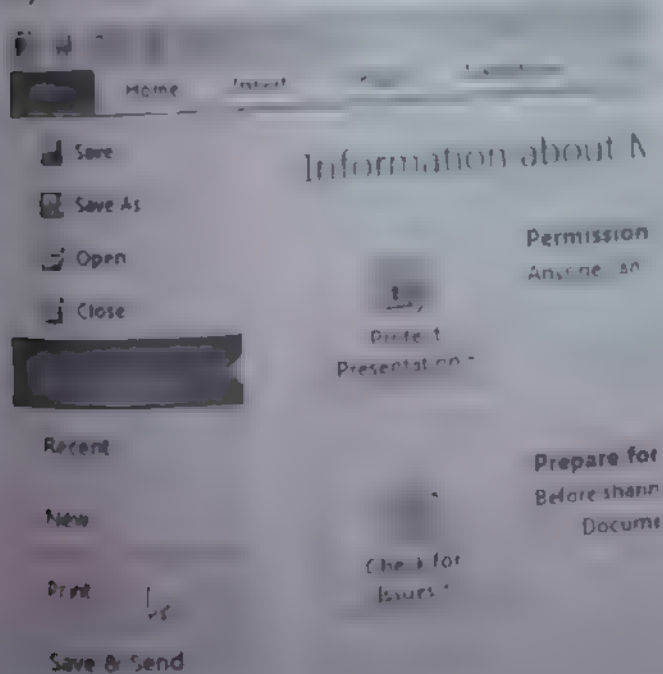


Figure 7.31 Printing slides

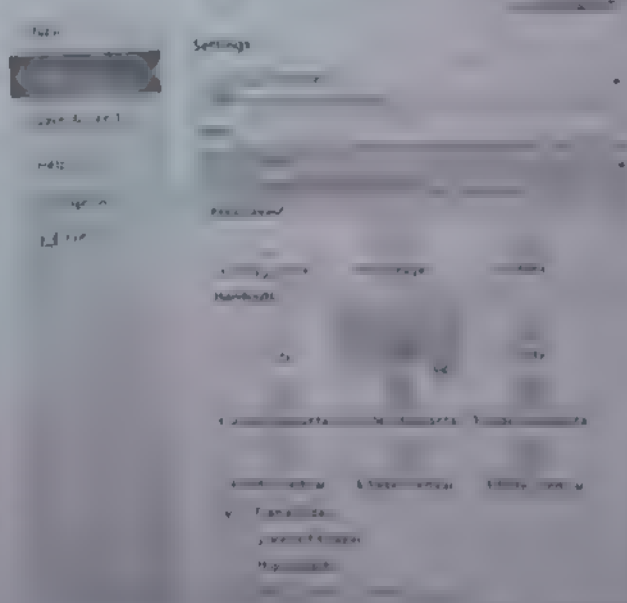


Figure 7.32 Printing handouts

7.7 SLIDE SHOW

Once your presentation is complete, you will now have to present it to an audience. PowerPoint has several tools and features to make your presentation smooth, engaging, and professional. In this section we will learn how to play a slide show and advance through slides. In addition, we will see how to set transition and slide timings and automate our presentations.

7.7.1 Running a Slide Show

PowerPoint runs your slide show from the first slide or from any slide of the slide show. Once the slide show has started, you can decide how to advance through the slides. Follow the steps given below to run a slide show:

Step 1: Click on the Slide Show tab.

Step 2: From the Slide Show group, click on the From Beginning command (as shown in Fig. 7.33) to start the slide show with the first slide. To start the slide show from any other slide, select the slide and click on From Current Slide. This option is useful if you want to view only certain slides.

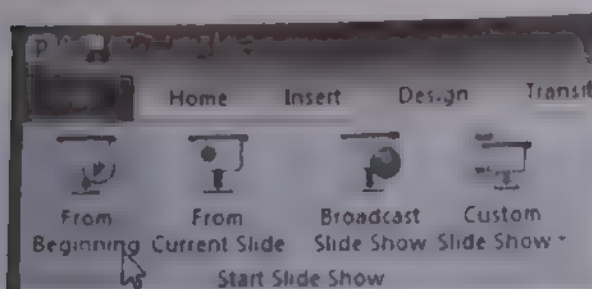


Figure 7.33 Starting presentation from the first slide



Figure 7.34 Slide Show view

An alternate way to start the slide show is to select Slide Show view at the bottom of the window as shown in Fig. 7.34.

Click on the right arrow to advance slides and the left arrow to reverse slides. These arrows are present on the bottom-left corner of the screen. You can also use the arrow keys on your keyboard to advance and reverse slides.

To stop or end a slide show, right click on the box and select End Show as shown in Fig. 7.35. An alternate way to end the presentation is to press the Escape key.

7.7.2 Setting the Timing and Speed of a Transition

Till now you have seen that in slide show, you advance to the next slide by clicking the mouse or pressing Enter key on the keyboard. If you want your presentation to run automatically, then you will have to set time after which the next slide will appear automatically. This feature is especially useful for untended presentations, such as at a tradeshow booth or an exhibition. Follow the steps given below to set the timing and speed of slide transition so that your presentation can run automatically:

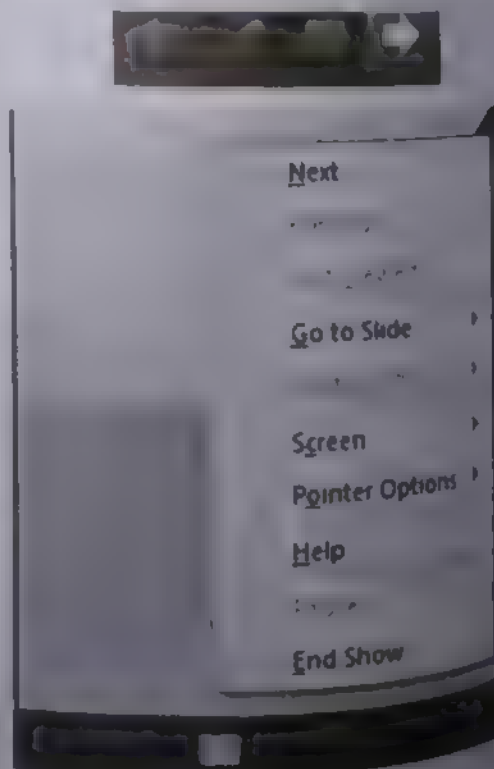


Figure 7.35 End show

- Step 1:** Select the slide.
- Step 2:** Click on the Transition tab.
- Step 3:** From the Timing group, under Advance Slide, uncheck the box next to On Mouse Click (this is used to advance the slide when the mouse is clicked) (as shown in Fig. 7.36).
- Step 4:** In the After field, enter the amount of time for which you want the slide to be displayed.



Click on Apply to All if you want all the slides to be transitioned after the same time.

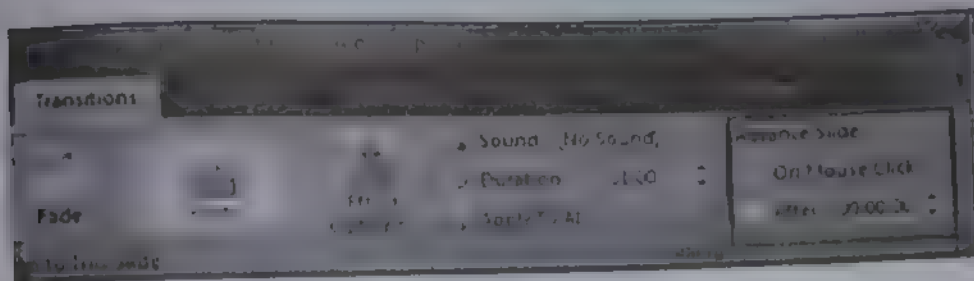


Figure 7.36 Setting the speed and timing of transitions effects

You can also control slide transitions effects.

- Step 1:** In either the Slide Sorter or Normal view, select the slide(s) on which you want to apply the transition.
- Step 2:** Click on Transition tab
- Step 3:** The most commonly used animations are displayed in the center. Click on the down arrow to see more effects. You can see the preview when you hover the mouse over an effect.
- Step 4:** Click on any transition from the list as shown in Fig. 7.37
- Step 5:** In the Duration option, enter the speed at which you want to play the transition.



Figure 7.37 Controlling transitions

- Step 6:** In the Sound field, click on the drop down menu to select a sound to play during a slide transition (optional).
- Step 7:** Select on Apply to All to apply the same transition to all your slides.

7.7.3 Automating a Slide Show

To fully automate a slide show you need to use the Rehearse and record slide timings feature of PowerPoint 2013

- Step 1:** Click on the Slide Show tab.
- Step 2:** From the Set Up group, click on Rehearse Timings

Step 3: The Rehearsal toolbar as shown in Fig. 7.38 will be displayed on the screen. It will automatically start timing the presentation. While timing your presentation, you can perform one or more actions of the following on the Rehearsal toolbar:

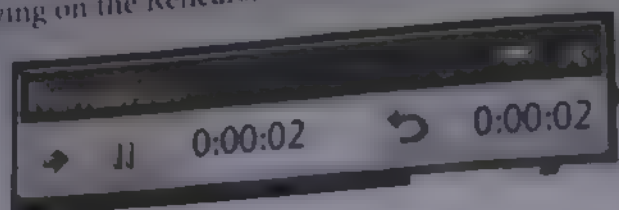


Figure 7.38 Rehearsal timing toolbar

- (a) Click on Next to move to the next slide.
- (b) Click on Pause to temporarily stop recording the time.
- (c) Click on Pause to restart recording the time after pausing.
- (d) Type the length of time in the Slide Time box to set an exact length of time for a slide to be displayed on the screen.
- (e) Click on Repeat to restart recording the time for the current slide.

After you set the time for the last slide, a message box displays the total time for the presentation and asks you to keep the recorded slide timings. Click on Yes to record and click on No to discard the recorded slide timings.

Finally, Slide Sorter view will be opened to display the time of each slide in your presentation.

To automate a slide show, you can also record narration in it. A recorded PowerPoint presentation can be played back at any time. The prerequisite of recording narration is to have a microphone.



A recorded narration increases the size of presentation. Therefore, such a presentation takes more time to download.

Follow the steps given below to record narration in your PowerPoint presentation:

Step 1: Click on Slide Show.

Step 2: From the Set Up group, click on Record Slide Show.

Step 3: The Record Slide Show box appears as shown in Fig. 7.39. Check the second option. The first option can be checked if you have recorded timing for each slide.

Step 4: Click on Start Recording to begin recording. The slide show will begin. Speak into the microphone to record your voice. Click to advance.

When the narration for the final slide is done, right-click on the slide and select End Show. You will be asked to save the timings. Click on Yes and the narration will be completed.

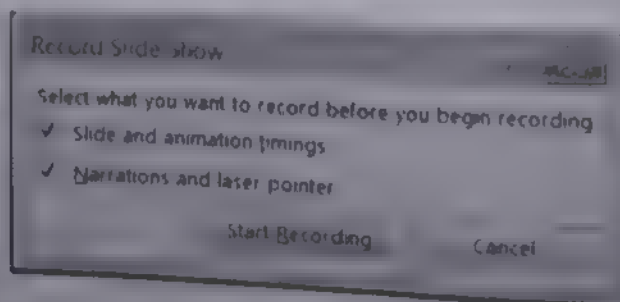


Figure 7.39 Record Slide Show box

Summary

- Microsoft PowerPoint is a presentation software developed by Microsoft.
- Slide Views are a group of four buttons present on the left of the Zoom control, towards the bottom of the screen. It helps you to switch between PowerPoint views.
- To make your presentation contain thoughtful arrangement of elements and colour, fonts, effects,

style, and layout, you can create the presentation using a template

- PowerPoint allows users to create presentations containing text, images, sounds, and other media objects.
- Headers and/or footers are added to include important information about the slides

Glossary

Presentation Document prepared in Microsoft Presentation.

Slide area Area where the slide is created and edited.

Multiple-choice Questions

- PowerPoint has been developed by _____.
 (a) Microsoft (c) Google
 (b) Apple (d) Apache
- Presentation is saved with an extension _____.
 (a) .ppt (c) .pps
 (b) .pptx (d) .pptd
- _____ tab is used to create a new presentation.
 (a) File (c) Slide Show
 (b) Insert (d) Animations
- Tabs are present in the _____.
 (a) File (c) Ribbon
 (b) Insert (d) Animations
- _____ is the area where the slide is created and edited.
 (a) Title bar (c) Tabs
 (b) Ribbon (d) Slide area
- _____ displays a list of topics which you can browse to find details about a topic of your choice.
 (a) Status bar (c) Help
 (b) Search bar (d) Ribbon
- _____ consists of a slider that can be slid left or right to zoom in or out
 (a) Status bar (c) View buttons
 (b) Zoom control (d) Title bar
- _____ view is useful while editing individual slides and rearranging them.
 (a) Normal layout (c) Reading
 (b) Slide sorter (d) Notes section
- _____ is not displayed on the screen during the presentation.
 (a) Normal layout (c) Reading
 (b) Slide sorter (d) Notes section
- Undo, Redo, and Save buttons are found on _____.
 (a) Status bar
 (b) Quick Access Toolbar
 (c) Title bar
 (d) Ribbon
- _____ displays all the thumbnails of all the slides in sequence
 (a) Slide tab
 (b) File tab
 (c) Insert tab
 (d) Animations tab

12. To apply 3D effects on a shape you must click the _____ tab
 (a) Slide (c) Insert
 (b) File (d) Format
13. WordArt feature is present in the _____ tab
 (a) Slide (c) Insert
 (b) File (d) Format
14. To end a presentation press the _____ key.
 (a) Esc (c) Alt
 (b) Shift (d) End
15. Rehearse Timing command is available in the _____ tab
 (a) Slide Show (c) Insert
 (b) File (d) Format
16. _____ view displays the time of each slide in your presentation.
 (a) Normal Layout (c) Notes Section
 (b) Slide Sorter (d) Reading
17. To record narration in your presentation, you must have a _____.
 (a) Speaker (c) Microphone
 (b) Webcam (d) Headsets
18. Which type of file cannot be inserted in a presentation?
 (a) .avi (c) .docx
 (b) .wav (d) None of these
19. Motion path is the _____.
 (a) transition between slides
 (b) entrance animation
 (c) exit animation
 (d) directing of the movement of an element on a slide
20. To give a uniform look to your presentation, you will use _____.
 (a) a background picture
 (b) a background color
 (c) a template
 (d) an animation
21. To add times to the slides in a presentation, you will use the _____.
 (a) Slide Show menu
 (b) Rehearse timing button
 (c) Slide transition button
 (d) Slide Sorter
22. _____ is not a view in PowerPoint
 (a) Slide Sorter (c) Slide Show
 (b) Normal (d) Notes
23. The best view to apply transition effects to multiple slides together is the _____ view.
 (a) Slide Sorter (c) Slide Show
 (b) Normal (d) Notes
24. Format painter is used to _____.
 (a) paint pictures on slides
 (b) copy formatting from one object to another
 (c) change the background color of slides
 (d) change the font style, size, and color of a text
25. The design template is applied to _____.
 (a) the current slide only
 (b) all the slides
 (c) every presentation you create
 (d) all of these
26. A new presentation can be created using a the _____.
 (a) blank presentation
 (b) existing presentation
 (c) design template
 (d) all of these
27. Which combination of keys inserts a new slide in the current presentation?
 (a) Ctrl + N (c) Ctrl + S
 (b) Ctrl + M (d) Ctrl + P
28. Which command is used to go to the first slide in a presentation?
 (a) Next slide button (c) Ctrl + Home
 (b) Page up (d) Home
29. _____ view shows thumbnails of all slides in a presentation.
 (a) Slide Show (c) Notes Page
 (b) Slide Sorter (d) Reading
30. The entry effect as one slide replaces another during a slide show is called _____.
 (a) animation (c) motion path
 (b) slide transition (d) slide navigation
31. _____ provides a printed copy of your presentation.
 (a) Notes

- b) Thumbnails
 c) Audience Handouts
 d) Presentation
- c) The _____ key will not move to the next slide during slide show.
 a) Esc (c) Enter
 b) Spacebar (d) Right arrow
- d) _____ is the default page orientation for notes pages, outlines, and handouts.
 a) Vertical (c) Portrait
 (b) Landscape (d) Horizontal
34. _____ is the default page orientation for printing slides in the presentation
 (a) Vertical (c) Portrait
 (b) Landscape (d) Horizontal
35. _____ changes the fill color of an object back to the default color.
 (a) Template (c) Patterns
 (b) Automatic (d) Fill colors

Answers to Multiple-choice Questions

- | | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1. (a) | 2. (b) | 3. (a) | 4. (c) | 5. (d) | 6. (c) | 7. (b) | 8. (a) | 9. (d) | 10. (b) |
| 11. (a) | 12. (d) | 13. (c) | 14. (a) | 15. (a) | 16. (b) | 17. (c) | 18. (d) | 19. (d) | 20. (c) |
| 21. (b) | 22. (d) | 23. (a) | 24. (b) | 25. (b) | 26. (d) | 27. (a) | 28. (c) | 29. (b) | 30. (b) |
| 31. (c) | 32. (a) | 33. (c) | 34. (b) | 35. (b) | | | | | |

Spreadsheet

Syllabus Mapping	Unit
Managing worksheets: formatting, entering data, editing, and printing a worksheet; handling operators in formula, project involving multiple spreadsheets, organizing charts and graphs, pivot table Spreadsheet functions: mathematical [SUMIF, SQRT, SUBTOTAL, SUMPRODUCT etc.], statistical [AVERAGE, STDEV, VAR, CORRELATION, REGRESSION, etc.], financial [PMT, RATE, PV, FV, NPV, IRR, NPV, Data Table, etc.], logical [AND, OR, IF, etc.]; date and time, lookup and reference, database and text functions. Creating spreadsheet in the area of: loan and lease statement; ratio analysis, payroll statements; capital budgeting; depreciation accounting; graphical representation of data, frequency distribution and its statistical parameters; correlation and regression.	Module II Unit 8

8.1 INTRODUCTION

A spreadsheet is a table of values arranged in rows and columns. Each value in the table has a pre-defined relationship with other values in the table. Therefore, if you change one value, the other value may also get affected.

Spreadsheet applications, also known as spreadsheets, are programs that help users create and manipulate spreadsheets electronically. They are specifically designed to ease the management of numbers and calculations.

MS Excel is a commercial spreadsheet application developed by Microsoft. It can perform calculations, and analyse and integrate data from different programs. It is used to calculate totals, averages, percentages, budgets, and complex financial and scientific formulas. Besides these functions, MS Excel is used for database management and doing business analysis. By default, documents saved in Excel 2013 are saved with the .xlsx extension.

In this chapter we will explore the most widely used spreadsheet software—MS Excel—in detail. The chapter will cover the basic elements of an electronic spreadsheet; teach you to create, open, save, and print it, enter text, number, date, formulas, and functions in the cells; insert or delete rows and columns, change the height and width of cells; represent information using charts; sort and filter data for analysis, and create macros and install add-ins to further increase your productivity.

8.2 ELEMENTS OF ELECTRONIC SPREADSHEET

In this section, we will read about the different elements of MS Excel. But before that, let us learn how to open the program.

Step 1: Click Start button.

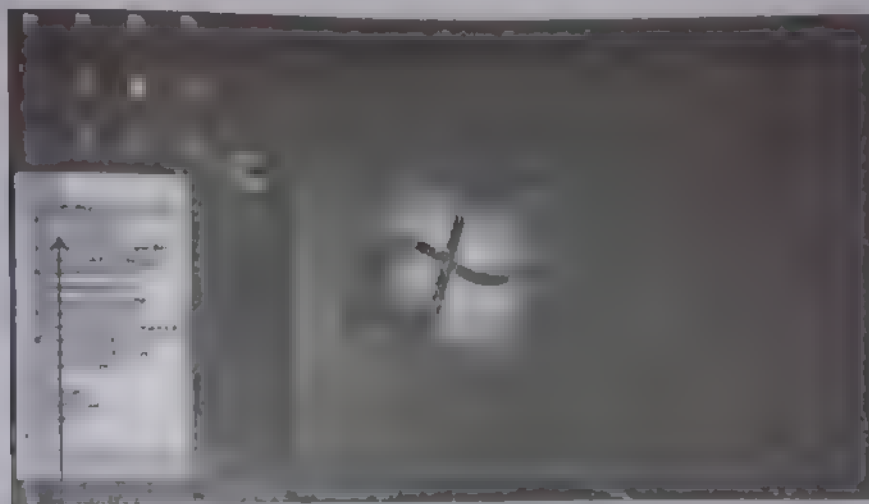
Step 2: Click on All Programs from the Start menu.

Step 3: Click on MS Office.

Step 4: Click on MS Excel 2013 as shown in Fig 8.1

When MS Excel starts, the Excel window will be displayed as shown in Fig 8.2. Let us now read about the elements in this window. A page in Excel is called a worksheet or simply sheet.

File Tab This tab is used to create new sheets; open, save, and print them; and do other file-related operations.



Microsoft Excel 2010

Figure 8.1 Starting MS Excel

Quick Access Toolbar Like in Word, this toolbar has buttons for most frequently used commands. You can customize this toolbar based on your choice.

Ribbon The Ribbon (as shown in Fig 8.3) contains commands organized in tabs and groups.

Tabs Tabs can be seen across the top of the Ribbon. They contain groups of related commands. Some tabs are Home, Insert, Page Layout, Formulas, Review, etc.

Groups Related commands are combined in groups. For example, commands like Cut, Copy, Paste, etc. are all placed in the Clipboard group.

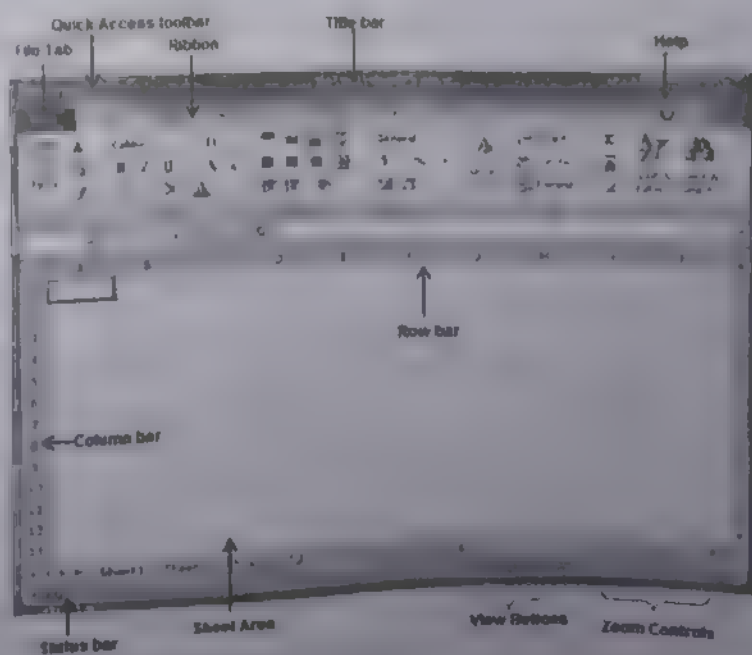


Figure 8.2 MS Excel window

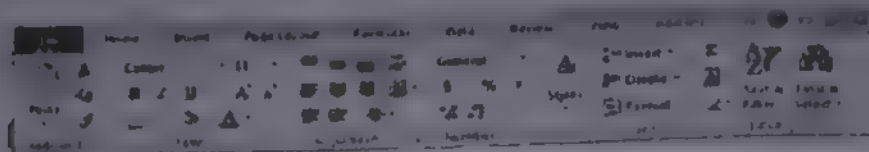


Figure 8.3 Ribbon

Title Bar The Title Bar appears next to the Quick Access Toolbar. It displays the name of the worksheet and the name of the program (which is MS Excel).

Help The Help icon is used to get help regarding the features of Excel.

Zoom Control It helps you to get a closer look at the text. It has a slider that can be dragged towards the left or right direction to zoom in or out. You can click on the – and + buttons to increase or decrease the zoom factor.

View Buttons The group of three buttons located to the left of the Zoom control, near the bottom of the screen, is used to switch among different sheet views. There are three views in Excel.

- **Normal Layout view** displays page in normal view.
- **Page Layout view** gives a full screen look of the sheet and displays pages as they will appear when printed.
- **Page Break view** shows a preview of where pages will break when printed.

Sheet Area Data is entered in the Sheet Area. The flashing vertical bar also known as the insertion point represents the location where text will be entered while typing.

Row Bar Rows are numbered from 1. They keep on increasing as you enter data. The maximum number of rows in Excel is 1,048,576.

Column Bar Columns are numbered starting from A. They keep on increasing as you enter data. After Z, it will start series of AA, AB and so on. You can have a maximum of 16,384 columns.

Status Bar It displays information like insertion point location, total number of pages and words in the document, etc. Status bar can be customized by right-clicking anywhere on it and by selecting or deselecting options from the provided list.

8.2.1 Opening a Spreadsheet

To open a workbook, follow the steps given below:

Step 1: Click on File tab.

Step 2: Click on Open.

Step 3: Select the folder in which the file to be opened is stored.

Step 4: Click on the name of the file.

Step 5: Click on Open.

OR, if you want to open a sheet that had been used recently, simply go to File tab and click on Recent as shown in Fig. 8.4. If the name of the file appears in the list of Recent files simply click on it to open it.

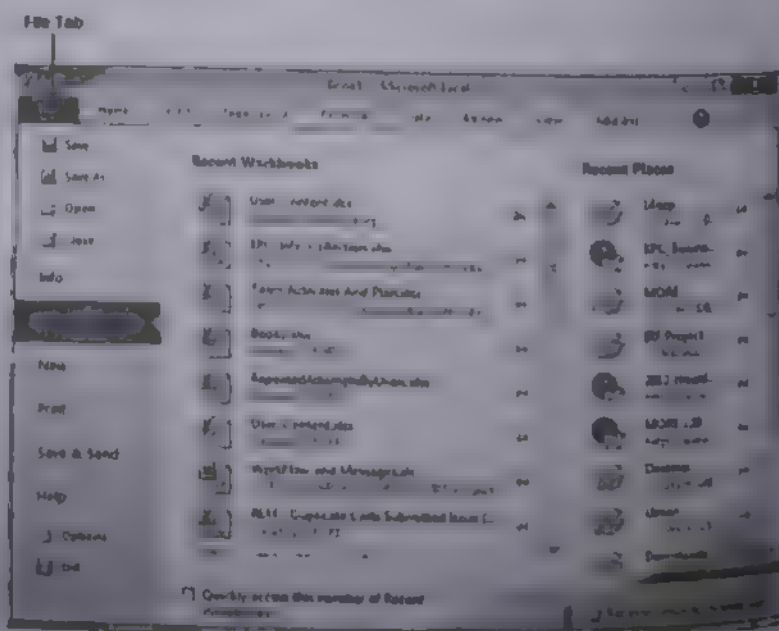


Figure 8.4 Opening workbook

8.2.2 Addressing Cells

In Excel, a cell is addressed using a combination of its row number and column number. When writing the address of a cell the column letter always precedes the row number. For example, in Fig 8.5, the address of the selected cell is C8.

8.2.3 Printing Spreadsheets

To print a sheet, follow the steps given below:

Step 1: Click on File tab.

Step 2: Click on Print.

Step 3: Click on Print button as shown in Fig. 8.6.

OR you can also press Ctrl + P and then click on the Print button (or press Enter).

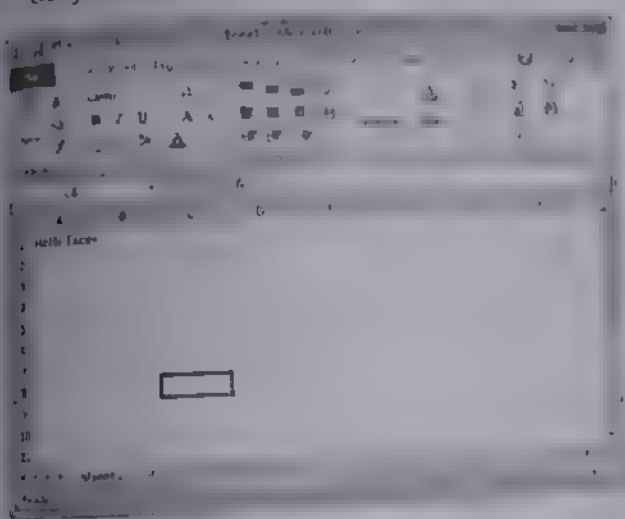


Figure 8.5 Cell addressing

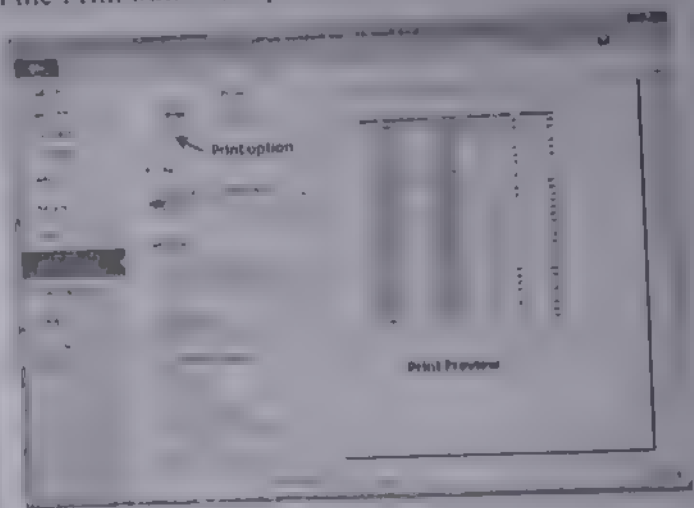


Figure 8.6 Printing a worksheet

However, before printing a sheet you can also choose a printer of your choice and even select which parts of the document to print. For example

- Active Sheets option is used to print the active sheet or sheets that you selected.
- Entire Workbook option prints the entire workbook, including chart sheets.
- Selection option lets you print only the range the sheets have been selected for

8.2.4 Saving Workbooks

When you open Excel, you get a workbook that has three worksheets. Once you are done with typing in the new Excel sheet, you must save your workbook to store it for future use. Saving a workbook automatically saves all its worksheets.

Therefore, to save a workbook, follow the steps given below:

Step 1: Click on File tab.

Step 2: Click on Save As option.

Step 3: Select a folder where you want to save your workbook. Enter the name with which you want to save as shown in Fig 8.7

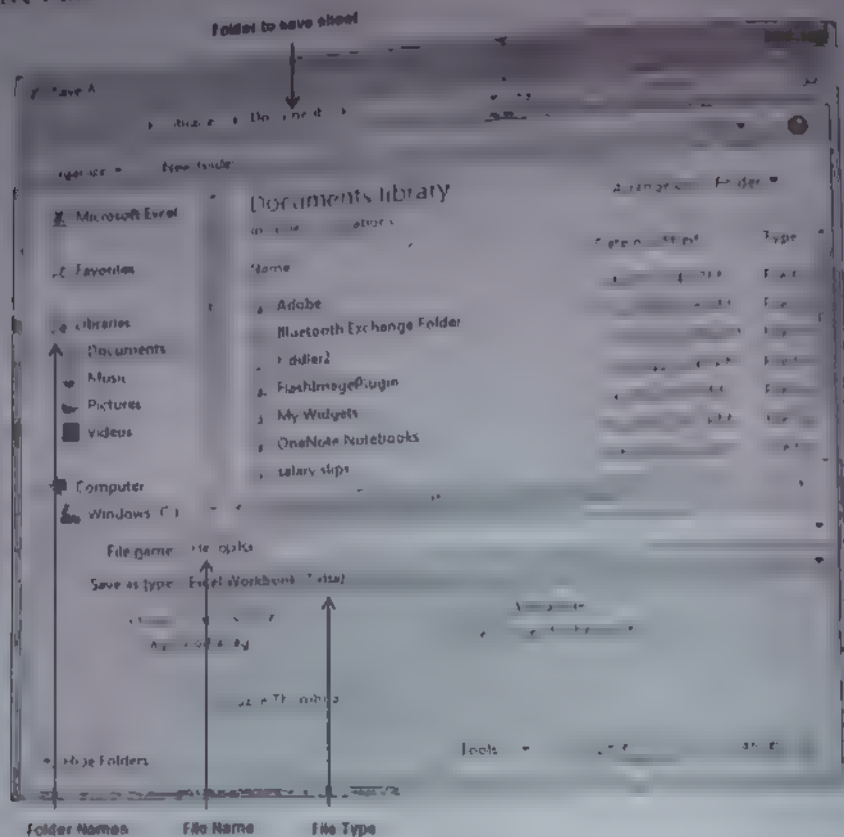
Step 4: Click on Save button

Figure 8.7 Saving worksheet

However, if you are editing an already saved sheet, then you must regularly save your work so that changes are made permanently in the worksheet. To save your work under the same file name, you can use any one of the three methods

- Press Ctrl + S keys.
- Click on the Save icon in the Quick Access Toolbar
- Click on File tab and then click on Save option.

Saving Excel File as .csv and .pdf

Many a times, we want to save an Excel file with an extension other than .xls. This means that you want to save your Excel file in some other file format.

The pdf file format makes the file as a portable electronic document that can be easily shared, especially via the web and e-mail. The csv file format is used to take the data (in a spreadsheet) and put it in a format that can be easily imported to other databases and applications. CSV file

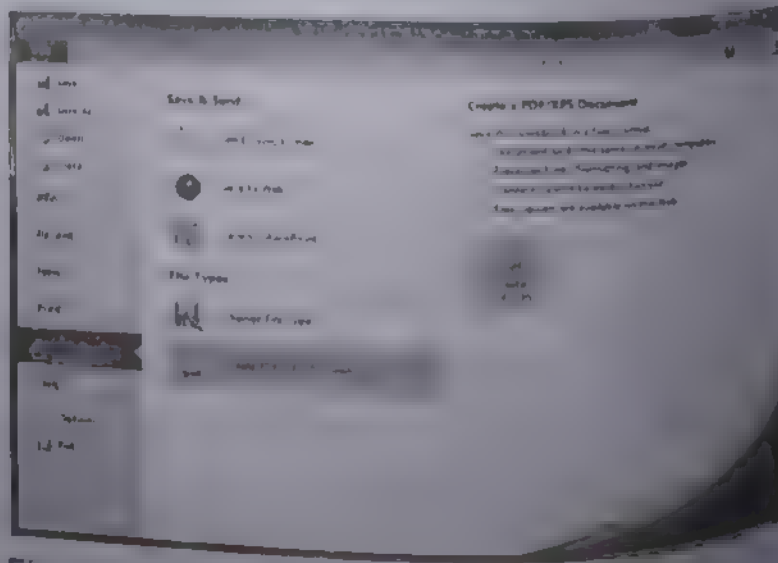


Figure 8.8 Saving file as PDF

can be easily read by a large number of database systems. Once the Excel file is saved as .csv, it can be used to perform more complex reporting functions. Follow the steps given below to save your Excel file as a PDF or CSV file.

Step 1: Open the workbook you want to save in another format.

Step 2: Click on File tab.

Step 3: Click on Save & Send.

Step 4: Create PDF/XPS Document.

Step 5: Click on Create PDF/XPS as shown in Fig. 8.8.

To save the file with .csv extension, Click on File tab. Select Save As option. In the Save As type box, select the option with .csv.



When you save your file in another format, some of its formatting, data, or features may not be saved.

8.3 MANIPULATION OF CELLS

We know that a cell is an individual addressable unit in an Excel sheet. It is formed by the intersection of a row and a column. In this section, we will learn about entering data, editing it, and changing the height and width of the cells.

8.3.1 Entering Text, Numbers, and Dates

To enter data in an Excel sheet, follow the steps given below.

Step 1: Activate the cell as shown in Fig. 8.9

Step 2: Type text, numbers, or date.

Step 3: Press enter or any Arrow key.

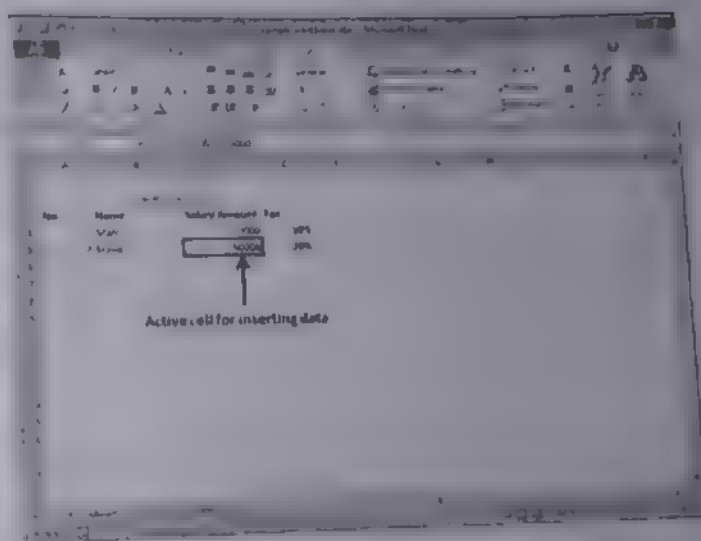


Figure 8.9 Activating cell



Press Tab to go to next column. Press Enter to go to next row.

An MS Excel cell can hold data of different types like numbers, currency, dates, etc. You can set the cell to hold these different types of data in the following way:

Step 1: Right click on the cell.

Step 2: From the pop-menu that appears, click on Format cells.

Step 3: Click on Number.

Step 4: From the various cell formats, select the desired type of data and its format. The various cell formats are as follows.

- General is default cell format of Cell.
 - Number displays cell as number with separator, as shown in Fig. 8.10.
 - Currency displays cell as a currency value along with currency sign.
 - Accounting is similar to currency but is used for accounting purpose.
 - Date stores data in different formats of date as in 17-09-2013, 17th-Sep-2013, etc.
 - Time stores data as time in different formats as in 1.30PM, 13.30, etc.
 - Percentage displays cell as percentage with decimal places like 50.00%.
 - Fraction displays cell as fraction like 1/4, 1/2, etc.
 - Scientific displays cell as an exponential value like 5.6E+01.
 - Text displays cell as a normal text.
 - Special stores data in special formats like Zip code, Phone Number, etc.
 - Custom allows you to create a custom number format that is added to the list of number format codes. In Excel 200-250 custom number formats can be added.
- Move the mouse over different options, to see how the data will be formatted.

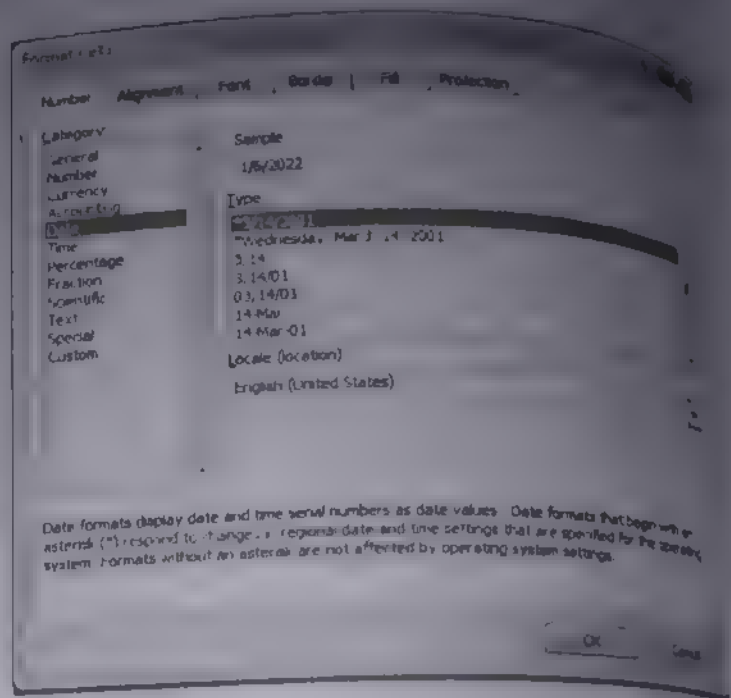


Figure 8.10 Entering numbers, text, or dates

8.3.2 Creating Text, Number, Date, Time, Weekdays, Months, or Years Series

Excel allows users to quickly fill cells with a series of texts, numbers, dates, times, weekdays, months, or years. For example, you can enter January in a cell, and then fill the cells below or to the right with February, March, April, etc.

Step 1: Select the cell(s) that contain(s) the starting number, date, time, weekday, month, or year.

Step 2: Drag the fill handle over the cells that you want to fill with the series as shown in Fig. 8.11.

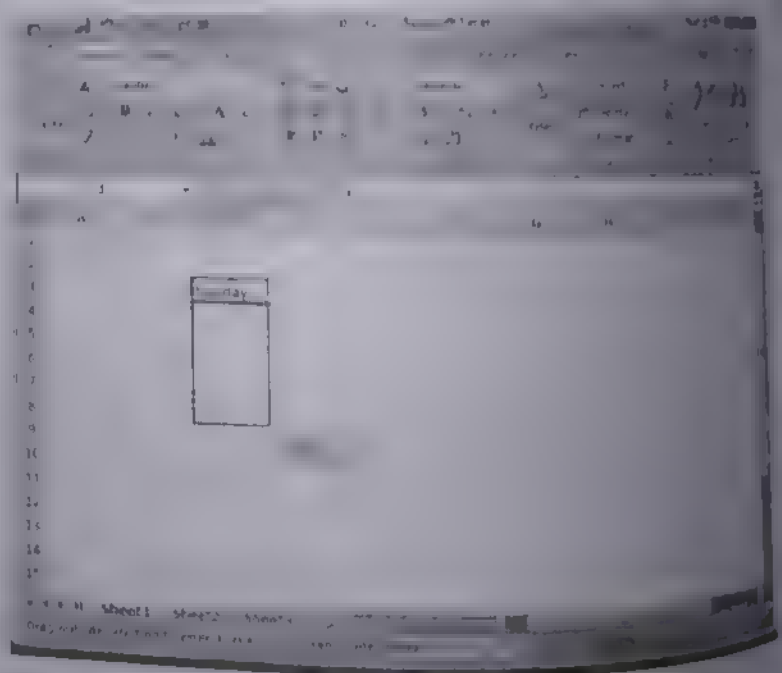


Figure 8.11 Creating series

8.3.3 Editing Worksheet Data

To modify data in a cell, follow the steps given below:

Step 1: Activate the cell.

Step 2: Enter a new value.

Step 1: Press enter or use any Arrow key to see changes.



A cell reference or address can also be specified as SheetName!Cell Address. For example, a cell B9 in Sheet2 can be referred as Sheet2!B9.

Cut, Copy, and Paste in MS Excel 2013

You can use the Cut, Copy, and Paste commands to move or copy data in an Excel sheet. To move data, follow the steps given below:

Step 1: Select the cell(s) which contain(s) data that you want to move.

Step 2: Click on the Home tab.

Step 3: Click on Cut command from the Clipboard group.

Step 4: Click on the cell where you want to paste (or move) the data.

Step 5: Click on the Home tab and then click on the Paste command from the Clipboard group



You can also press **Ctrl + X** to cut and **Ctrl + V** to paste. Also you can right-click on the cell and select Cut or Paste option from the menu.

To create a duplicate copy of data in a cell, perform the same steps but use the Copy command instead of the Cut command. Alternatively, you can press the **Ctrl + C** command to copy the data. To make multiple copies of data, press Copy as many times as you want the copies to be of.

Using Paste Special Options

When you paste data, MS Excel copies every piece of information in the range of cells you have selected. The Paste Special command offers much more options. To use the Paste Special command, click the down arrow below the Paste button in the Home tab. Then, select Paste Special option (refer Fig. 8.12) from the menu to open the Paste Special dialog box. Let's try to understand the meanings of all the options.

All Pastes everything in the selected cell including formulas, formatting, etc.

Formulas Pastes all the text, numbers, and formulas without any formatting.

Values Converts formulas in the selected cell to their calculated values.

Formats Pastes only the formatting from the selected cell and no values or formulas.

Comments Pastes only the notes that are attached with the currently selected cells.

Validation Pastes only the data validation rules.

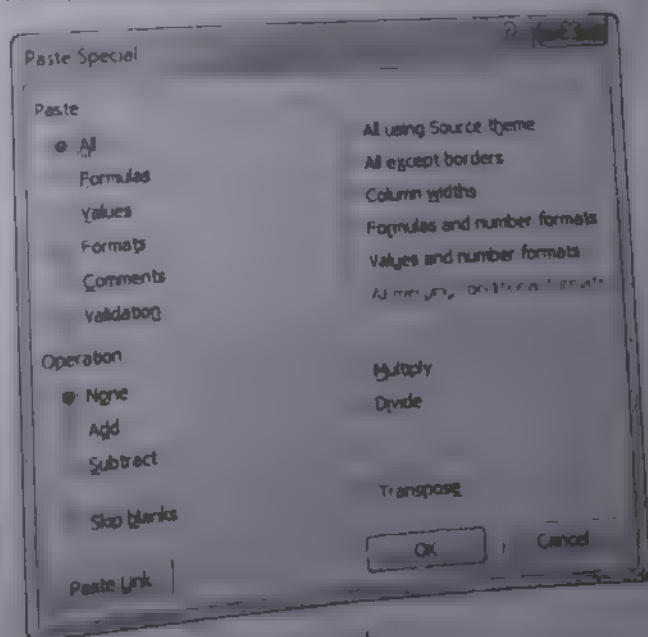


Figure 8.12 Paste special

All Using Source Theme Pastes all the information plus the cell styles.

All Except Borders Pastes everything without copying any borders.

Column Widths Applies column width of the cells copied to the clipboard to the columns where the cells are pasted

Formulas and Number Formats Assigns number formats to the pasted values and formulas

Values and Number Formats Converts formulas to their calculated values and assigns number formats to all the copied or cut values

All Merging Conditional Formats Pastes conditional formatting into the cell range.

While performing the Paste operation you can also perform simple math calculations based on the value(s) in the copied or cut cell(s) and the value in the target cell(s). Options given below cater to these needs

None No operation is performed. This is the default setting.

Add/Subtract/Multiply/Divide The values you cut or copy to the Clipboard are added/subtracted, multiplied/divided to the values in the cell range where you paste

Other options are as follows.

Skip Blanks This option is selected when you want to paste only from the cells that are not empty.

Transpose This option is selected when you want to change the orientation of the pasted entries. That is, data in rows are copied into columns and data in columns are copied into rows.

Paste Link This option is selected to create links between data

8.3.4 Inserting New Worksheet

MS Excel provides you three new blank sheets. In case you need more, you can insert a blank worksheet by following the steps given below:

Step 1: Right click on the Sheet Name and select Insert option as shown in Fig. 8.13.

Step 2: A dialog box opens. Select Worksheet option and click OK

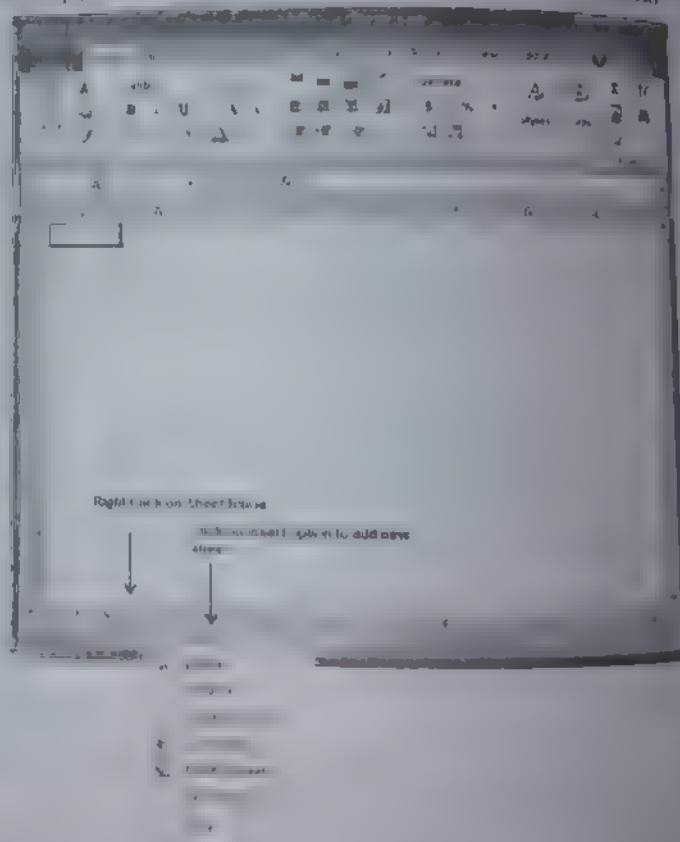


Figure 8.13 Inserting Worksheet



You can also press Shift+F11 keys for inserting a new sheet.

8.3.5 Inserting or Deleting Cells, Rows, and Columns

MS Excel allows you to insert blank rows/columns in the worksheet. To add a new row, follow the steps given below.

Step 1: Select the row above which you want to insert the new row. To insert multiple rows, select the same number of rows as you want to insert. For example, to insert five new rows, select five rows.

Step 2: Click on Home tab.

Step 3: In the Cells group, click the triangle next to Insert option as shown in Fig 8.14.

Step 4: Click on Insert Sheet Rows.

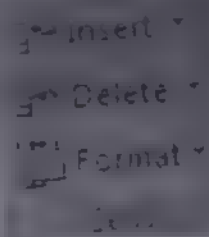


Figure 8.14 Cells group



You can also right-click on the selected row and click on Insert option from the pop-up menu.

When a new row is inserted, all row numbers are adjusted accordingly wherever they are referenced. To insert a new column, follow the same steps. Select the column, click on Home tab, and select Insert Sheet Columns. You can also select a column, right-click on it, and select Insert to add a new column. Whenever you add a column, column numbers are adjusted accordingly wherever they are referenced.

To delete rows/columns from a sheet, follow the steps given below.

Step 1: Select rows or columns to be deleted.

Step 2: Click on Home Tab.

Step 3: From the Cells group, click on the triangle next to Delete option.

Step 4: Click on delete rows/columns.

Whenever a row/column is deleted, the row/column numbers are adjusted accordingly wherever they are referenced.

8.3.6 Changing Column Width and Row Height

On a worksheet, you can specify a column width of 0-255 characters. This means that you can display a maximum of 255 characters in a cell that is formatted with the standard font. The default column width is 8.43 characters. If you set the column width to 0, the column is hidden.

Similarly, you can specify a row height of 0-409 points where 1 point equals approximately 1/72 inch. The default row height is 12.75 point. If you set the row height to 0, the column is hidden. To change the column width or row height, follow the steps given below.

Step 1: Select the column(s).

Step 2: Click on Home tab.

Step 3: In the Cells group, click on Format.

Step 4: Click on Row Height and enter the new height. Click on Column Width and enter the new width.

Step 5: Click on OK to apply the changes.

You can also change the width of column(s) by dragging the right boundary of the column heading until the column has the desired width.

Similarly, to change the height of row(s) drag the boundary below the row heading until the row has the desired height as shown in Fig 8.15.

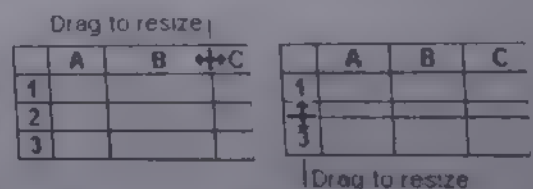


Figure 8.15 Changing column width and row height

8.3.7 Formatting Rows, Columns, and Sheets

Most of the commands to format text are present in the Home tab under Font, Alignment, and Number groups. Font commands help you to change the style, size, colour of text; apply borders; and fill colours to cells. Alignment commands help you to format how the text is displayed across cells horizontally and vertically. Number commands help you to change how cells display numbers and dates. Follow the steps given below to format the text with different options in the Fonts group:

Step 1: Select the cells and click on the Home tab.

Step 2: Click on the down arrow next to the Font textbox.

Step 3: Move your mouse over the various fonts to see a preview of the fonts.

Step 4: Click on any particular font of your choice as shown in Fig. 8.16.

Step 5: To change the font size, click on the down arrow next to the font size button.

Step 6: Move the mouse over different sizes to see a preview of how the text will look.

Step 7: Click on the font size of your choice as shown in Fig. 8.17.

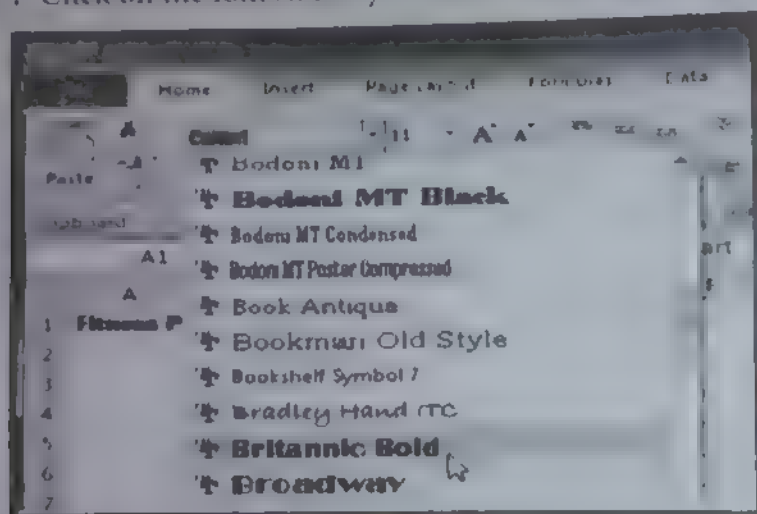


Figure 8.16 Font type

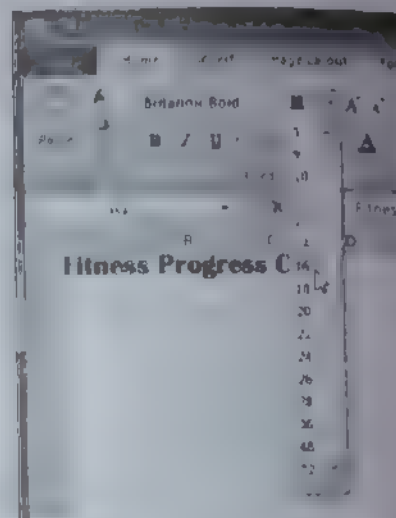


Figure 8.17 Font size

Step 8: To change the font colour, click on the down arrow next to the font colour button.

Step 9: Move the mouse over different colours to see a preview of how the text will look.

Step 10: Click on the colour of your choice as shown in Fig. 8.18.

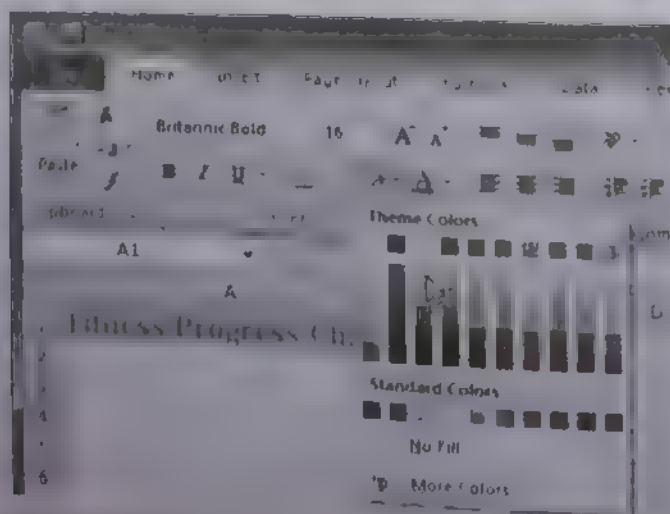


Figure 8.18 Font colour

Step 11: To apply Bold, Italics, and Underline effects (refer Fig. 8.19) to your text, select the text and press corresponding commands from the Font group.

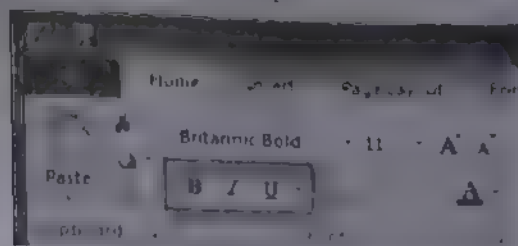


Figure 8.19 Bold Underline, Italics

Step 12: To apply border to selected cell(s), click on the down arrow next to the Borders command and select a style that you want. You can even change the colour and style of border as shown in the figure.

Step 13: To add colour, select the cells and click on the drop down arrow next to Fill colour command.

Step 14: Move the cursor over the various Fill colours to see a preview in the worksheet.

Step 15: Click on the colour of your choice as shown in Fig. 8.20.

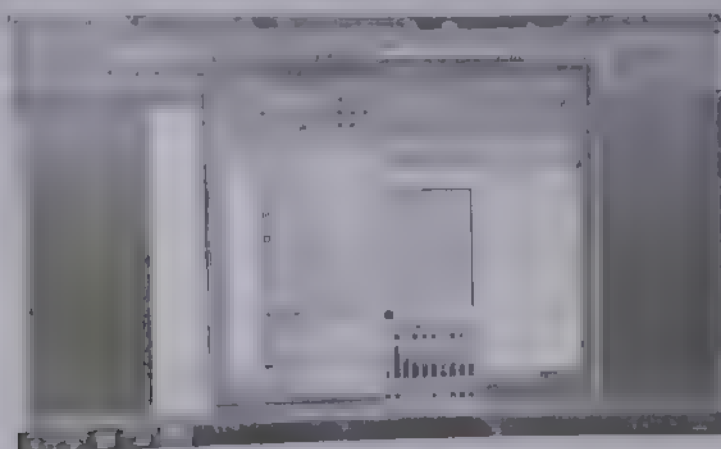


Figure 8.20 Fill colour

Using the Formatting Toolbar

To format the cells with the toolbar, select the cells and then click a button to apply one of the formats shown in Table 8.1

Table 8.1 Formatting options

		Effect
\$	Currency	Displays dollar sign: 75.30 as \$75.30
%	Percent	Displays a number as a percentage: 45 as 45%
,	Comma	Same as currency, but without dollar sign: 12345.6 as 12,345.60
→ 0 0 0	Increase Decimal	Displays one more place after the decimal: 0.45 as 0.450
0 0 → 0	Decrease Decimal	Displays one place less after the decimal: 0.450 as 0.45

Using the Format Cells Dialog Box

Excel has more numeric formats which you can select from the Format Cells dialog box. Select the cells you want to format.

Then open the dialog box by selecting Format _ Cells from the Menu bar. The Format Cells dialog box has separate pages for Number, Alignment, Font, Border, Patterns, and Protection as shown in Figure 8.21.

Aligning Text

By default, Excel left aligns texts and right aligns numbers. You can use the buttons on the Formatting toolbar to override the defaults and align the text and numbers as required within the cells, such as left, center, and right:

Left

Center

Right

The following default settings reflect some standard rules for aligning texts and numbers:

1. Columns of the text should be left aligned
2. Columns of the numbers should be kept in the default (right) alignment and formatted so that the decimal points align
3. Column labels should appear over the contents of the column. If the column is filled with numbers, then the heading should be right aligned or centered. Labels for text columns should be left aligned or centered.

Merge, Shrink to Fit, and Wrap Text

1. Select the title and several additional cells below the title
2. Select Format _ Cells and check one or all of the following options in the Format Cells dialog box
 - Merge Cells check box merges the cells.
 - Shrink to Fit option reduces the size of the content within the selected cells so that the contents get fit.
 - Wrap Text option wraps the content of a cell if it exceeds the cell's boundaries.

How to Format Numbers

Excel lets you present numbers in a variety of formats. Formatting is used to identify the numbers as currency or percentages and to make numbers easier to read by aligning decimal points in a column. You can format the selected cells with the following three tools:

1. Formatting toolbar
2. Format Cells dialog box
3. Shortcut menu

When you format a number, you change its appearance and not its numeric value. The default format for numbers, that is, General, does not display zeroes, and this does not have any effect on the value of the number. For example, if you enter 10.50, only 10.5 would be displayed as it has the same numeric value. Thus, Excel does not display the extra or trailing zero.

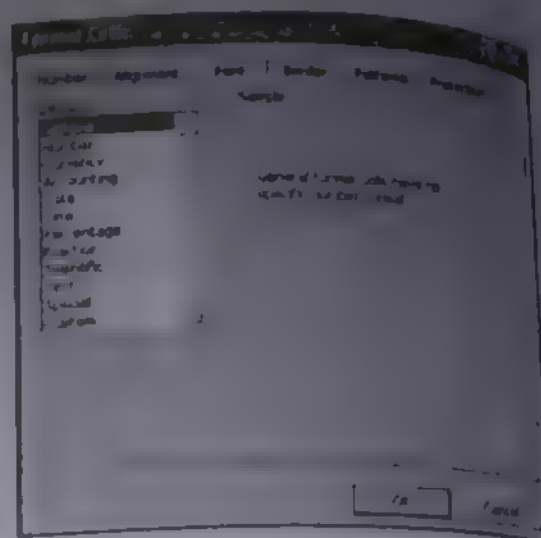


Figure 8.21 Format Cells dialog box

Linking Data in Rows, Columns, and Sheets

MS Excel has a powerful feature of linking data. When the data of a cell(s) is linked to other cells, the data is dynamically pulled from the source cell(s) into the linked cell(s). Cell(s) may be in the same workbook or in different workbooks.

The feature of linking data eliminates the need to enter same data in multiple sheets. This not only saves time but also reduces errors and improves data integrity. The moment the source data is changed, all the sheet(s) linked to the source data will also be updated automatically to reflect that change.

For example, a company can store the prices of all its products in one worksheet. It can then link data with all other worksheets that need product pricing details. To link data in rows, columns or sheets, follow the steps given below.

Step 1: In the source worksheet, select the cell(s) you want to link.

Step 2: Click on the Home tab and click on copy command.

Step 3: Move to the destination worksheet and click on the cell(s) you want to link.

Step 4: Click on the Home tab and click on the down arrow below the Paste command.

Step 5: Click on Paste Link as shown in Fig. 8.22.

Step 6: Return to the source worksheet and press the Esc key to remove the border around the selected cell(s).

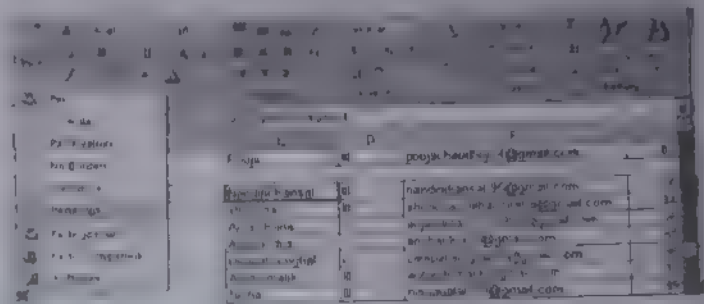


Figure 8.22 Paste Link



After creating the link between data of two different workbooks, if you move any of the workbook in a different folder, the link may be broken.

8.3.8 Protecting and Unprotecting Cells, Rows, Columns, and Sheets with or without Password

Before protecting data in cell(s) first un-protect all the cells in your worksheet. To do this, follow the steps given below.

Step 1: Select all of the rows and columns in your sheet.

Step 2: Right click on them, and select Format Cells from the pop-up menu.

Step 3: Click on Protection and uncheck the Locked checkbox as shown in Fig. 8.23.

Step 4: Click on OK.

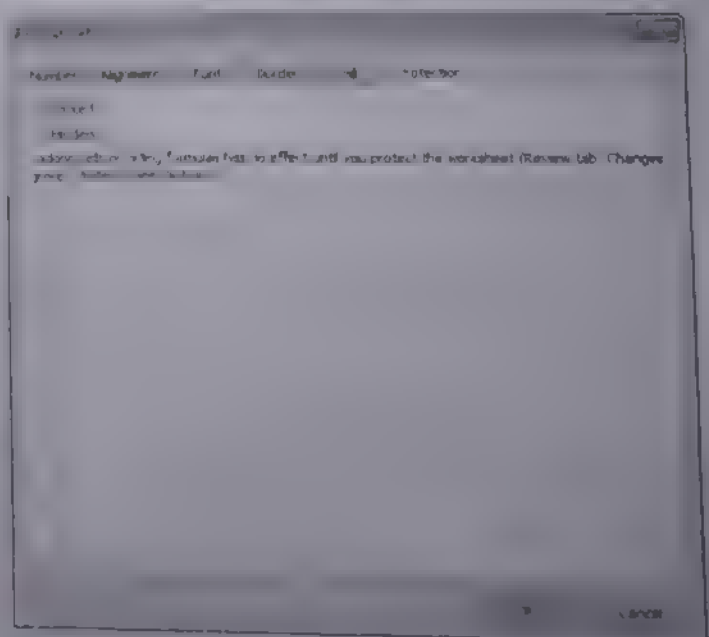


Figure 8.23 Protection tab

Now, to protect your rows and columns perform the steps given below:

Step 1: Select the cell(s) you want to protect.

Step 2: Right click on them and select Format Cells from the pop up menu.

Step 3: Click on Protection tab and check the Locked checkbox.

Step 4: Click on OK.

Step 5: For locking cells you must also protect the worksheet. So click on the Review tab from the ribbon.

Step 6: Click on the Protect Sheet button as shown in Fig. 8.24.

Step 7: In the Protect Sheet window, you may enter a password to protect the sheet (optional) as shown in Fig. 8.25.

Step 8: Click on OK.

Figure 8.24 Protecting sheets

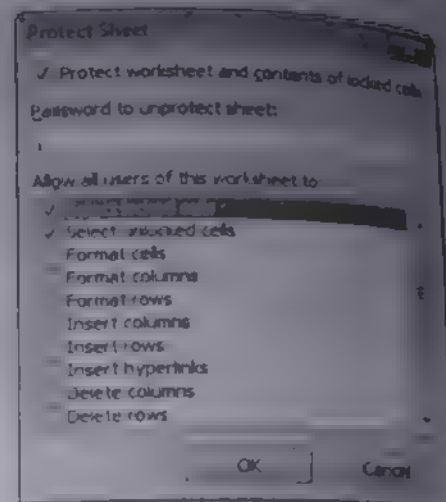


Figure 8.25 Setting passwords

8.3.9 Hiding/Unhiding Rows, Columns, and Sheets

Excel 2013 has an excellent feature to hide the selected rows, columns, and sheets to protect specific content from any sort of detrimental usage. To hide rows, follow the steps given below:

Step 1: Select the row(s) you want to hide.

Step 2: Click on the Home tab.

Step 3: Click on the down arrow next to Format command in the Cells group.

Step 4: Move the mouse over Hide & Unhide option. Select Hide Rows from the menu as shown in Fig. 8.26.

The selected rows will be hidden.

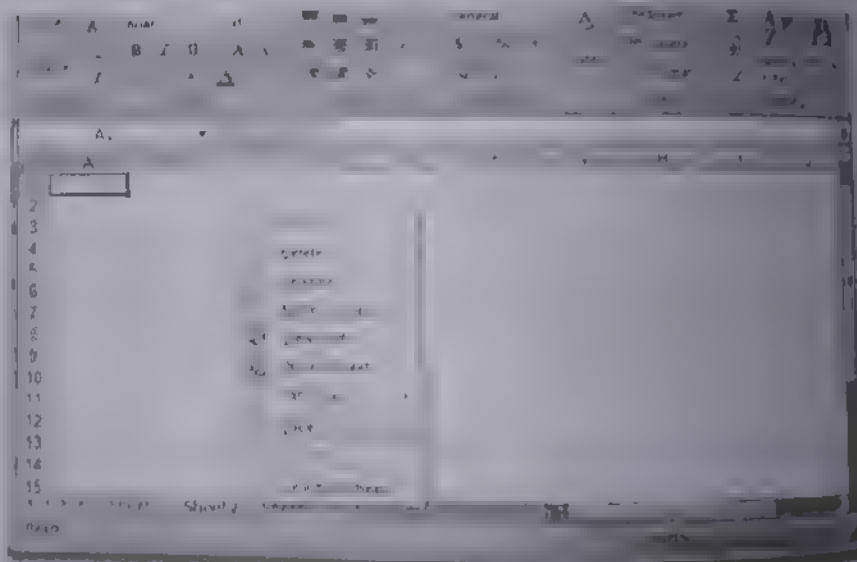


Figure 8.26 Hiding sheets



To select multiple consecutive rows click on the rows while holding the Shift key. To select multiple rows that are not consecutive, click on the rows while keeping the Ctrl key pressed.

To hide a sheet, right-click on the sheet. From the menu that appears click on Hide. Alternately, select the slide. Then, from the Cells group in Home tab, click on the down arrow next to Format command and click on Hide & Unhide. Select Hide Sheet option.

Follow the steps given below to unhide a row, column, or sheet.

Step 1: Select the row(s), column(s), or sheet(s) you want to unhide.

Step 2: Click on Home tab.

Step 3: From the Cells group, click on the arrow next to Format command and click on Hide & Unhide as shown in Fig. 8.27.

Step 4: From the menu, click on Unhide Row/Column/Sheet option as shown in Fig. 8.28.

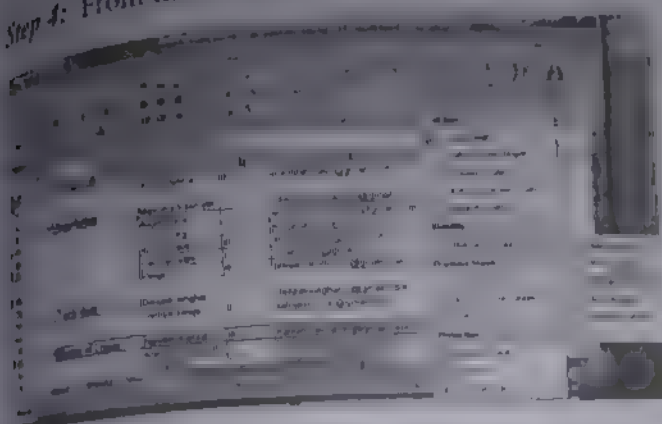


Figure 8.27 Hide and unhide option

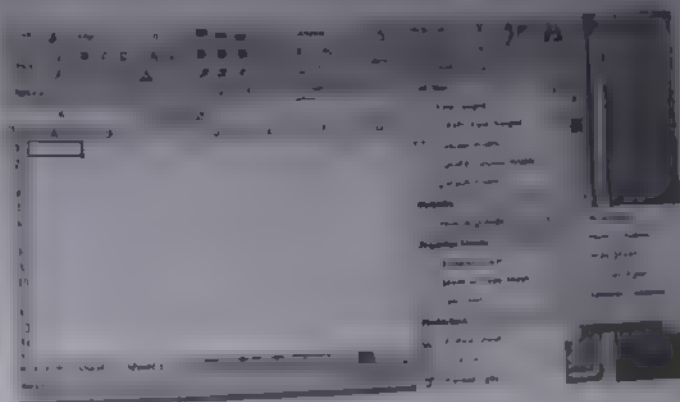


Figure 8.28 Unhide row/column

8.4 FUNCTIONS AND CHARTS

The best feature of Excel is that it allows users to analyse data. For this, it has a number of functions. Many formulas that you create can use the available functions. By using these functions, you can perform calculations that are difficult if you use only the operators. For example, the LOG or SIN function used to calculate the Logarithm or Sin ratio is difficult to be specified using just operators. In this section, we will learn about these exciting features of MS Excel.

8.4.1 Formulas

Formulas are the bread and butter of worksheet. Without them, worksheet will have just a simple tabular representation of data.

A formula is an expression that performs a calculation. One of the most exciting features of Excel is that it can calculate expressions using a cell address. The cell's address represents the value in a cell. This is called using a cell reference.

The main advantage of using formulas is that if value(s) change(s) in the cell(s), formulas automatically calculate the updated results. A formula must begin with the $=$ sign and can have mathematical operators, such as $+$, $-$, $*$, $^$ or logical operators (like and, or).

The result of formula is displayed in the cell that has been currently selected and contains the formula.



When you click on a cell that has a formula, the formula will be displayed in the formula bar.

In Excel, every formula must begin with an equal to sign ($=$). The steps for entering formulas in an Excel sheet are as follows:

Step 1: Select the cell where the result has to be displayed.

Step 2: Type the $=$ sign.

Step 3: Type the cell address that contains the first number in the equation.

Step 4: Type the operator you need in the formula as shown in Fig. 8.29.

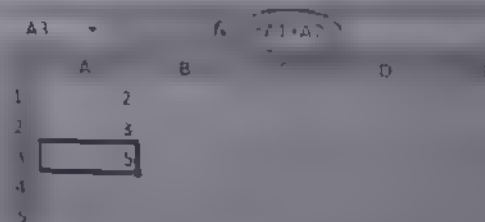


Figure 8.29 Entering a formula

Step 5: Type the cell address that contains the second number in the equation.

Step 6: Press Enter.

The expression will be calculated, and the result will be displayed in the cell. For example, you can write $A1 + A2$ to add the values in cell A1 and A2.

If by mistake, you have typed the wrong formula and you want to edit it, then follow the steps given below

Step 1: Select the cell that contains the formula.

Step 2: The formula will appear in the Formula bar. Click in the formula bar and change the formula.

Step 3: Press the Enter key.

Can you guess the meaning of $= A1 < C12$? Yes, it compares the value of cell A1 with that of C12. If the two values are same, the result is TRUE; otherwise FALSE.

The formula specified in one cell can be copied and applied to other cells. This can be done by simply clicking on the cell having the formula and then dragging the mouse across other cells. This will copy the formula in all the cells.

While dragging, Excel automatically updates the cell references. This means that MS Excel adjusts the cell references in the original formula to suit the position of the copies that you make. This is called *relative cell addresses*.

Types of Cell References

The term *cell reference* means the cell to which another cell refers. For example, if in cell A1 you write $= A2$, then A1 has the same value as that of A2, meaning it refers to A2. In Excel, there are three types of cell references

Relative The cell address changes as you copy or move it. This is because the cell reference is relative to its location

Absolute The cell reference does not change even when you copy or move the cell to any other cell.

Mixed Either the row or the column changes (but not both) when you copy or move.

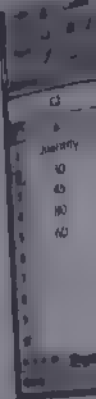
Example of Relative Cell References Suppose in cell C2 we specify a formula $= A2 * B2$ (refer Fig. 8.30a). If we copy the cell into cell C3, the formula gets copied and automatically the cell references are updated as shown in Fig. 8.30b. In C4, the formula bar shows $= A4 * B4$. Thus, we see that each new cell updates relative to the new location. Relative cell references makes it easy to copy cells as it avoids having to edit each cell to ensure pointing to the correct place.



You can even specify a range of cells by using colon between the top-left and bottom-right cells. For example, A1:C5 includes all the cells from A1 to C5

Example of Absolute Cell References In absolute reference, we use the \$ sign to specify cell reference. For example, if we write $= \$A\1 in any cell, \$ in front of column A means 'do not change the column' and \$ in front of row 1 means 'do not change the row' when a cell is moved or copied to any other cell (as shown in Fig. 8.31a). For example, if we write $= \$A\$2 * \$B\2 in cell C2 and then copy the contents of C2 in other cells, then we see in the figure that the cell reference does not change and each cell in column C has the same value (as shown in Fig. 8.31b).

Figure 8



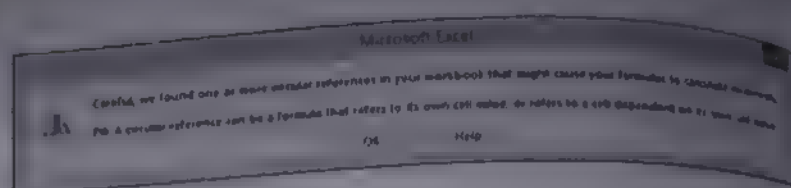
Figure

Example
the oth
will re
where

Circ
ence
it m
refer



(a)



(b)

Figure 8.33 Example of circular cell reference

References to Other Worksheets We know that a workbook can have multiple worksheets. To reference a cell in a different worksheet, we write `= worksheet!cell`, where `worksheet` is the name of the worksheet and `cell` is the cell address. For example, if we type in cell A4 of sheet1 `= Sheet2!C5` then the value of cell A4 of sheet A1 will be equal to the value of cell C5 of sheet2.

8.4.2 Functions

Functions are predefined formulas and are already available in Excel. Some frequently used functions in Excel are given below.

Mathematical and Statistical Functions

Count and Countif The Count function is used to count the number of cells that contain numbers. The Countif function, as shown in Fig. 8.34, is used to count the number of cells that meets the specified condition.

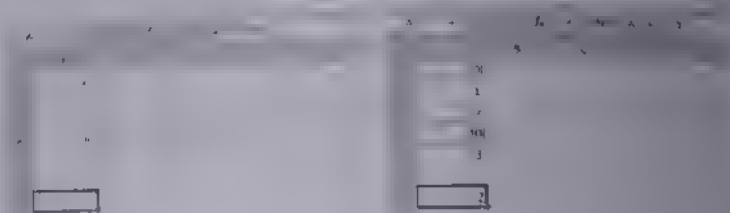


Figure 8.34 Count and Countif



The name of every function is followed by brackets. In between the brackets, the arguments (range of cells) are specified.

Countblank(range) This function finds the number of empty cells in the specified range. For example, `COUNTBLANK(A1:A5)` will find the empty cells from A1 to A5.

Sum and Sumif The Sum function is used to add values. The Sumif function, as shown in Fig. 8.35, sums the values, but only of those values that meet the specified condition.

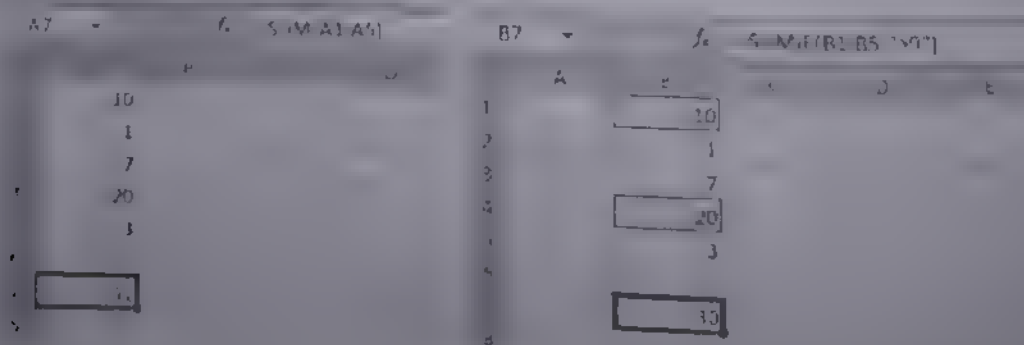


Figure 8.35 Sum and Sumif

Median, Mode, and Standard Deviation The Median function is used to find the median (or middle number). The Mode function is used to find the most frequently occurring number. Standard Deviation is used to find the standard deviation of values in a given range of cells. These functions are shown in Fig. 8.36.

f. =MEDIAN(A1:O1)							f. =MODE(A1:O1)							
A	B	C	D	E	F	G	A	B	C	D	E	F	G	
1	0	7	8	6	5	9	8	0	7	8	6	5	9	8
2														
3														
4														

Figure 8.36 Median and Mode

Average and Averageif The Average function is used to find the mean (or average) of values in a range of cells. Similarly, the Averageif function, as shown in Fig. 8.37, is used to find the average of only those values in a range of cells that meet a specified condition.

f. =AVERAGE(A1:O1)							f. =AVERAGEIF(A1:O1, >5)							
A	B	C	D	E	F	G	A	B	C	D	E	F	G	
1	0	7	8	6	5	9	8	0	7	8	6	5	9	8
2														
3														
4														

Figure 8.37 Average and Averageif

Min and Max Min is used to find the minimum value from a range of values and Max is used to find the maximum as shown in Fig. 8.38.

f. =MIN(A1:O1)							f. =MAX(A1:O1)							
A	B	C	D	E	F	G	A	B	C	D	E	F	G	
1	0	7	8	6	5	9	8	0	7	8	6	5	9	8
2														
3														
4														

Figure 8.38 Min and Max

Large and Small This function is used to find the largest and the smallest value as shown in Fig. 8.39. As a second argument to these functions you can specify a number- n to find the nth smallest or nth largest value.



Abs gives the absolute value. **Sign** gives the sign of a number (+ve, 0, -ve). **Sqrt** gives the square root of a number. **Mod** divides two numbers and gives their remainder.

f. =SMALL(A1:O1, 3)						
A	B	C	D	E	F	
0	7	8	6	5	9	

A	B	C	D	E	F
0	7	8	6	5	9

8

0

Figure 8.39 Large and small

Round This function rounds (Fig. 8.40) a number to the specified number of decimal digits

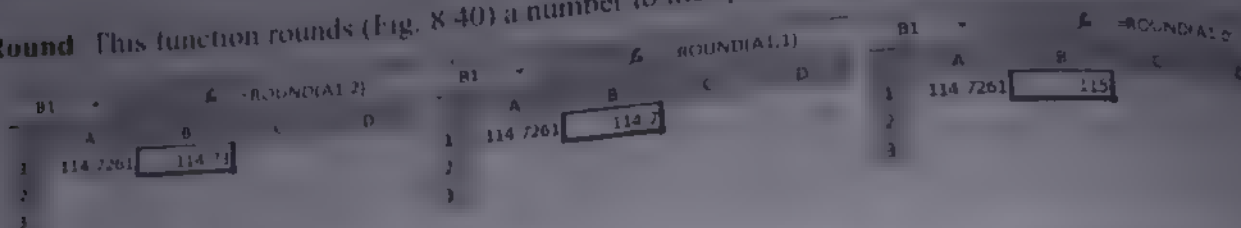


Figure 8.40 Round

Trunc(number, no of digits) The Trunc function truncates a number by removing the fractional part. For example, Trunc(123.34567, 2) returns 123.34. If you do not specify no_of_digits then the entire fractional part will be removed. For example, Trunc(34.567) returns 34.

Rand() This function generates a random number that lies between 0 and 1. For example Rand() may return 0.876239.

Ceiling(number, significance) This function rounds a number to the nearest multiple of significance. For example, Ceiling(34.25, 0.1) = 34.3, Ceiling(45.25, 0.5) = 45.5.

Convert(number, from unit, to unit) This function converts a number from one unit into another unit. For example, CONVERT(60, "m", "sec") returns 3600.

Floor(number, significance) This function rounds off the number down to the nearest integer towards zero or to the nearest multiple of significance. For example, Floor(156.32, 1) = 156, Floor(3.4, 2) = 2. The first number is the one that has to be rounded off and the second number is the multiple to which you need to round off. In the example Floor(3.4, 2), the nearest multiples are 2 and 4, but the one down towards zero is 2; hence, the answer.

INT(number) This function rounds down a number to the next lowest integer. For example, INT(123.45) = 123, INT(-8) = -9.

EVEN(number) This function returns a number next to the nearest even integer. For example, EVEN(231,34) = 232.

LCM(num1, num2, ...) This function returns the smallest number that can be divided by each of the given numbers. For example, LCM(8, 20) = 40.

Dollar(number, [decimal_places]) In this function the second argument is optional. The function converts a number into text using the currency format. For example, Dollar(9876, 2) = \$9,876.00.

AVEDEV(num1, num2, num3, ...) This function calculates the average deviation of a supplied set of events. For example, AVEDEV(10.7, 7.2) results in 1.65.

Forecast This function returns the prediction of a future value based on existing values.

Correlation

The correlation coefficient is a value between -1 and +1, which tells us how strongly two variables are related to each other. A correlation coefficient of +1 indicates a perfect positive correlation. For example, if one variable increases, the other also increases. Similarly, if one variable decreases, so does the other variable. To find the correlation between two variables we use the Correl function, which takes two variables as its input, as shown in Fig. 8.41.

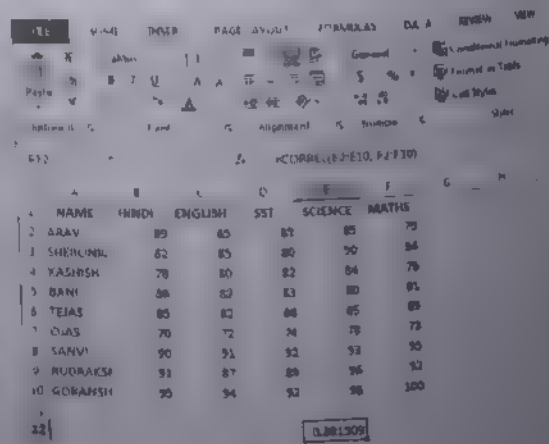


Figure 8.41 Correl

Regression Analysis

In statistics, regression analysis is done for estimating the relationships among variables. It is especially used when the focus is on the relationship between a dependent variable and one or more independent variables (or 'predictors'). Regression analysis helps us to understand how the typical value of the dependent variable changes when any one of the independent variables is varied, while other independent variables are held fixed. The following steps (Fig. 8.42) are used to do regression analysis:

Step 1: Select the data.

Step 2: Click on the Insert tab.

Step 3: Click on Scatter in the Charts area and then choose the first scatter graph.

Step 4: Right click on a data point on the graph and click on 'Add trendline' from the drop down menu.

Step 5: Scroll to the bottom of trendline options and click the check boxes for 'Display R-squared value on chart' and 'Display equation on chart'. R-squared and the regression equation will appear on the chart.

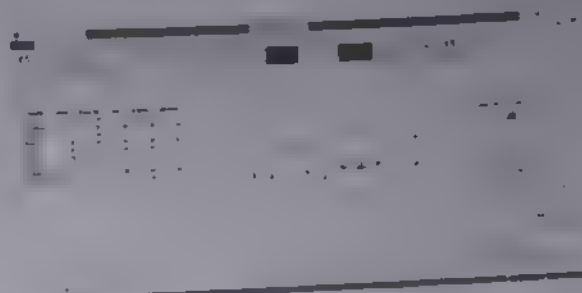


Figure 8.42 Regression analysis

VAR and VARA

The VARA() function calculates the sample variance of a supplied set of values. Its syntax is VARA(number1, [number2], ...). Where the number arguments are values that provide a minimum of 2 numeric values to the function. You can enter up to 255 values in this function.

The VARA() function in Excel is very similar to the VAR() function, in that both functions calculate the sample variance of a supplied set of values. However, the difference between them is that when an array of values, containing text or logical values is supplied to the function, the VAR() function ignores the text and logical values, whereas the VARA() function assigns the value 0 to text and the values 1 or 0 to logical values (see Fig. 8.43).

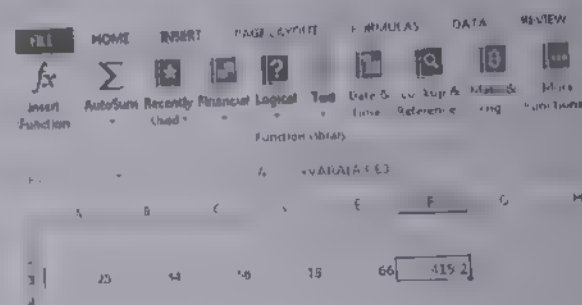


Figure 8.43 Vara

SUBTOTAL

The SUBTOTAL function returns the subtotal of the numbers in a column in a list or database. The syntax for the SUBTOTAL function is SUBTOTAL(method, range1, [range2, ..., rangen]) where, method can have any value from 1-11 or from 101-111. The functions corresponding to these numbers are given in Table 4.2. Basically, this argument specifies the type of subtotal to be created. Numbers from 1-11 denote that the function will have hidden values. However, those varying from 101-111 will ignore the hidden values in the calculation.

Table 8.2 SUBTOTAL function

Method	Function
1 or 101	Average
2 or 102	Count
3 or 103	Counta
4 or 104	Max
5 or 105	Min
6 or 106	Product
7 or 107	STDEV
8 or 108	STDEVP
9 or 109	SUM
10 or 110	VAR
11 or 111	VARP

range1, range2, ..., rangen specify the range of cells that need to be included in the subtotal.
Need for using SUBTOTAL. When there is a long list of records in a table and you have the grand total being displayed at the end, the SUBTOTAL function can be a big help. For example, look at Fig. 8.44 if we filter the Region column to display only WEST region's sales and we want the total to include only those items, then the SUBTOTAL function should be used. If you use the SUM function, the grand total of all the regions will be displayed.

The screenshot shows the Excel ribbon with the Home tab selected. The 'Filter' button is visible. Below the ribbon, a table is displayed with the following data:

	A	B	C	D	E
	ITEM	QTY	UNIT PRICE	PRICE	TOTAL
1					
5	D	WEST	25	290	7250
9	H	WEST	100	390	39000
13	L	WEST	45	375	16875
17	P	WEST	25	275	6875
21	T	WEST	110	560	61600

The formula bar shows the formula: `=SUBTOTAL(9, D5:D21)`. The status bar at the bottom indicates '5 OF 20 RECORDS FOUND'.

Figure 8.44 SUBTOTAL

To use the SUBTOTAL function, the following steps are to be used:

Step 1: Apply a filter to the list. In Fig. 8.44, we have used it on the Region column.

Step 2: Select the cell where you want the grand total.

Step 3: Click the AutoSum button on the Home tab (see Fig. 8.45). Since the list is filtered, a SUBTOTAL formula is inserted, instead of a SUM formula.

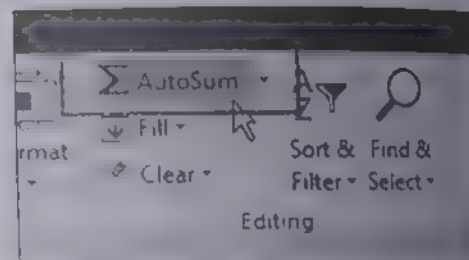


Figure 8.45 Autosum for subtotal

SUMPRODUCT

The SUMPRODUCT function multiplies the corresponding numbers in the given arrays, and then returns the sum of those products. The syntax of SUMPRODUCT function can be given as SUMPRODUCT(array1, [array2], [array3], ...), where, array1 is a mandatory argument that tells where in the sheet the first array can be found. The values in this array will be used to multiply with values in other arrays and then added; array2, array3, are optional arrays. You can specify maximum 255 arrays. Values of all the arrays will be used for multiplication and addition to produce the final result.

Points to remember

- The array arguments must have the same dimensions (i.e., the same number of rows and columns). If this condition is not satisfied, then SUMPRODUCT will return the #VALUE! error value.

The screenshot shows the Excel ribbon with the Home tab selected. The 'Filter' button is visible. Below the ribbon, a table is displayed with the following data:

	A	B	C	D	E
1	ARRAY 1		ARRAY 2		
2	1	2		3	4
3	5	6		7	8
4	9	10		11	12
5					
6	SUM OF PRODUCT =				

The formula bar shows the formula: `=SUMPRODUCT(A2:B4, D2:E4)`. The status bar at the bottom indicates 'Sheet1' and 'Sheet2'.

Figure 8.46 SUMPRODUCT example

- All values in the array must be numeric, otherwise SUMPRODUCT will treat them as zeros.
- If only one array is supplied, SUMPRODUCT will simply sum the items in the array.

Look at the example given in Fig. 8.46. If you write SUMPRODUCT(A2:B4, D2:E4) or SUMPRODUCT(A2:B4 + D2:E4) then, $(1+3) + (2+4) + (5+7) + (6+8) + (9+10) + (10+12)$ is done.

However, if you write, or SUMPRODUCT(A2:B4 * D2:E4), then, $(1*3) + (2*4) + (5*7) + (6*8) + (9*10) + (10*12)$ is done.

Similarly, SUMPRODUCT(A2:B4 * D2:E4) $(1/3) + (2/4) + (5/7) + (6/8) + (9/10) + (10/12)$. These examples are based on arrays arranged in columns

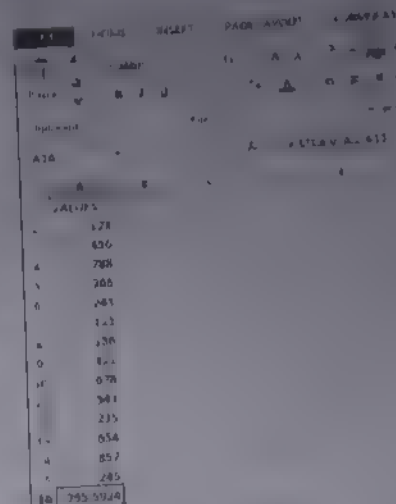


Figure 8.47 STDEV

STDEV

The STDEV function calculates standard deviation of a supplied set of values. The syntax to use this function is STDEV(number1, [number2], ...), where, number1, [number2], ... are the values which will be used to calculate the standard deviation (see Fig. 8.47). You can provide a maximum of 255 numeric values, arrays of values, or references to cells containing numbers. However, for the function to work properly, you must supply at least two numeric values.

Points to remember

- STDEV assumes that all the values are a sample of the population. If data represents the entire population, then use STDEVP instead of STDEV.
- Logical values and text representations of numbers are also counted.
- Arguments that are error values or text that cannot be translated into numbers cause errors.
- If you want to include logical values and text representations of numbers in a reference as part of the calculation, use the STDEV.A function.
- STDEV uses the following formula:

$$\sqrt{\frac{\sum (x - \bar{x})^2}{(n - 1)}}$$

Date and Time Functions

Year, Month, Day, Date In Excel, you can enter date as 2-3-2016 or 2/3/2016, where month precedes day. Therefore, the date means 3rd February 2016. Similarly, you can enter time as 5:00. This is shown in Fig. 8.48.

The Year function is used to extract the year from a date. The Month function is used to extract month from a date and Day function extracts the day from a date. You can use the Date function to add the number of days to a date.

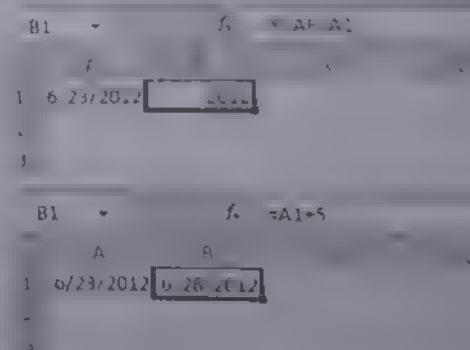


Figure 8.48 Year



Date function takes three arguments—year, month, and day. When days are added to a date, Excel automatically rolls it to the next month or year if need arises. For example, 12/27/2015 + 10 = 1/6/2016.

Now, Today, Time, Hour, Minute, and Second NOW() is used to get the current date and time. Similarly, TODAY() is used to get the current date. Hour() returns the hour, Minute() is used to get minutes, and Second() is used to get the seconds from current time. TIME() is used to add hours, minutes, and seconds to a time value. Excel automatically rolls time if required. For example, 23:15:00 + 2 = 1:15:00. This is shown in Fig. 8.49.

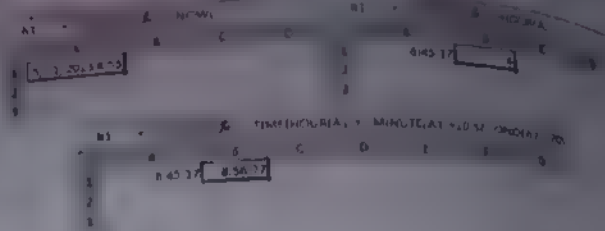


Figure 8.49 Now, Time, Hour, Minute, and Second

DAYS360 It returns the number of days between two dates based on a 360-day year (12 months of 30 days each). This is basically used in some accounting calculations like computing payments if your accounting system is based on twelve 30-day months.

The syntax of DAYS360 function is DAYS360(start_date, end_date, [method]). Note that the dates should be entered by using the Date function, or derived from the results of other formulas or functions. For example, use DATE(2016,1,9) to return 9th January 2016. Method is an optional that specifies whether to use the US or European method in the calculation.

YEARFRAC This function calculates the fraction of the year represented by the number of whole days between two dates (the start_date and the end_date). The syntax of this function is YEARFRAC(start_date, end_date, [basis]).

Here, basis is an optional argument that specifies which standards you want to follow—US or European, a 30-day year or 365-day year.

For example, (1/1/2016, 7/30/2016) gives the result 0.58055556.

WORKDAY This function returns indicated number of working days before or after a date (the starting date). Working days exclude weekends and any dates identified as holiday. The WORKDAY function is used to calculate invoice due dates, expected delivery times, or the number of days of work performed. The syntax of this function is WORKDAY(start_date, days, [holidays]).

Here, days specifies the number of non-weekend and nonholiday days before or after start_date. A positive value gives a future date and a negative value gives a past date.

And holidays is an optional argument that specifies one or more dates to exclude from the working calendar.

For example, workday(10/1/2016, 151) gives 212 as the result.



If start_date plus days gives an invalid date then an error is generated.

Text Functions

Join Strings This function is used to join two strings.

Left and Right Left function is used to extract leftmost characters from a string and Right function is used to extract rightmost characters as shown in Fig. 8.50.

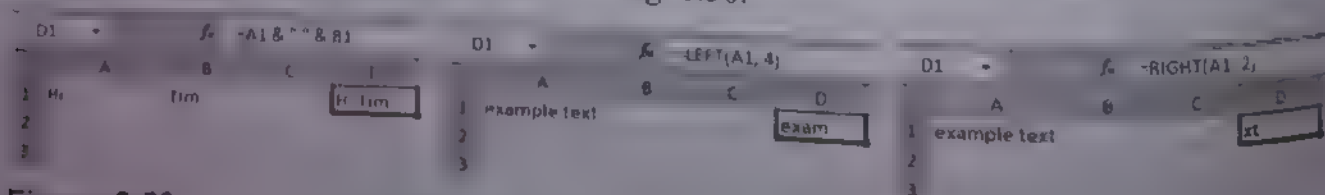


Figure 8.50 Join, Left, and Right

Mid and Len The Mid function is used to extract a substring. It accepts three arguments. First is the address of the cell, second is the starting position, and third is the number of characters to be extracted. Len function gives the length of the string as shown in Fig. 8.51.

	A	B	C	D
1	example text			=MID(A1, 5, 3) pie
2				
3				

	A	B	C	D
1	example text			=LEN(A1) 12
2				
3				

Figure 8.51 Mid and Len

Find and Substitute Find function is used to find the position of a substring in a string and Substitute function is used to replace a substring with another substring as shown in Fig. 8.52. While using the Find function, remember that it is case sensitive (characters in upper case are not same as those in lower case).



Concatenate: Combines two text strings. **Trim:** Removes spaces from left and right side of the text string. **Lower:** Converts all characters in the text string into lower case. **Upper:** Converts all characters of text string into upper case.

	A	B	C	D
1	example text			=FIND("pie", A1) 5
2				
3				

	A	B	C	D
1	example text			=SUBSTITUTE(A1, "pie", "tea") example tea
2				
3				

Figure 8.52 Find and Substitute

Logical Functions

Excel 2013 has a set of logical functions like AND, FALSE, IF, IFERROR, NOT, OR, and TRUE. All these functions return either TRUE or FALSE.

AND(arg1, arg2, ...) It returns TRUE if the conditions are true and FALSE otherwise. For example, AND(A1 >= 75, B1 >= 60) returns true only if both the conditions are true.

OR(arg1, arg2, ...) It returns TRUE if any one of the conditions is true and FALSE if none of the condition is true. For example, OR(A1 >= 75, B1 >= 60) returns true if either of the condition is true.

NOT(arg) It checks the argument. If it is true it returns false, and if the argument is false it returns true.

FALSE() It takes no argument and simply assigns FALSE in the cell.

TRUE() Like FALSE(), this function does not take any argument but simply assigns logical TRUE in the cell.

IF() The IF function (Fig. 8.53) is used to check whether a condition is true or not. It gives one (first) value if condition is true and another (second) value if it is false. To use this function, select the cell and type the function. You can also use this function along with AND/OR operator. If two or more conditions are specified in IF function, operator requires that all the conditions must be true to give true result. Similarly, the OR operator requires that at least one of the conditions must be true to give true result.

The syntax of IF statement is

IF (condition, value if true, value if false)

	A	B	C
1	75	=IF(A1 >= 75, 'Pass', 'Fail')	=IF(A1 < 75, 'Fail', 'Pass')
2		Pass	Fail
3			

Figure 8.53 IF function



The AND/OR function can check a maximum of 255 conditions.

Nested IF() The IF function can be nested, when there are multiple conditions to be met. In such a case the FALSE value is replaced by another IF function to test the additional condition as shown in Fig 8.54. You can nest multiple IF functions within one Excel formula. Excel allows you to nest up to 7 IF functions to create a complex IF THEN ELSE statement. The syntax of nested IF statement is

IF(condition1, value_if_true1, IF(condition2, value_if_true2, value_if_false2))

This is similar to writing an IF THEN ELSE statement:

```
IF condition1 THEN
    value_if_true1
ELSE IF condition2 THEN
    value_if_true2
ELSE
    value_if_false2
END IF
```

Figure 8.54 demonstrates the use of a nested if statement.

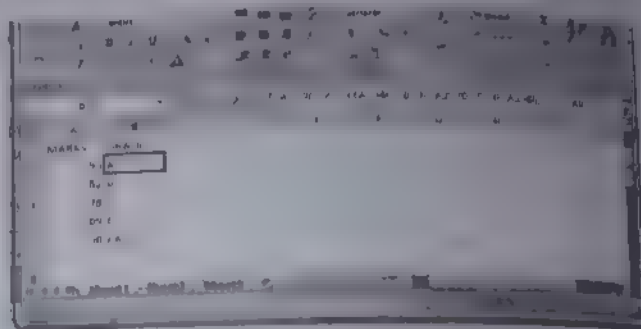


Figure 8.54 Nested IF function

Financial Functions

Microsoft Excel is a very important tool for investment bankers and financial analysts. They are masters in excel shortcuts and formulae. Though there are more than 50 financial functions in Excel, we will discuss some frequently used financial functions here. However, before we start, remember that if we pay an annual payment at an annual interest of 6% then for monthly payments, the rate will be $6\% / 12 = 0.5\%$.

Moreover, the number of periods (nper) is the number of times the payment is made until the entire amount is paid. For example, if we are making monthly payments for a 20 year loan, then $nper = 12 \times 20 = 240$.

Future value (FV) The FV function is used to find out the future value of a particular investment which has a constant interest rate and periodic payment. The syntax of FV is FV(Rate, Nper, [Pmt], PV, [Type]) where, Rate = interest rate, Nper = number of periods, [Pmt] = payment period, PV = present value, and [Type] = when the payment is made (if nothing is mentioned, then it means that the payment has been made at the end of the period).

Example Mittals have invested US\$1000 in 2015. The payment has been made yearly. The interest rate is 10% p.a. What would the FV be in 2019 (see Fig. 8.55)?

		=FV(C4,C5,C6,C7,C8)				
	A	B	C	D	E	F
1						
2						
3						
4		Rate	10%			
5		NPER	3			
6		PMT	1			
7		PV	-1000			
8		TYPE	0			
9						
10		FV				\$3,327.69
11						
12						

Figure 8.55 Future value

FVSCHEDULE The FVSCHEDULE function is used to calculate the future value with the variable interest rate. Its syntax is FVSCHEDULE (Principal, Schedule), where Principal = the present value

of a particular investment, and Schedule a series of interest rates put together (use different boxes and select the range).

Example Guptas have invested US\$1000 at the end of 2016. It is expected that the interest rate will change every year. In 2017, 2018, and 2019, the interest rates would be 7%, 9%, and 6% respectively. What would the FV be in 2019 (see Fig. 8.56)?

Present Value (PV) The Present Value function is used to find out PV. Its syntax is PV (Rate, Nper, Pmt, FV, [Type]), where, Rate = interest rate per period, Nper = number of periods, [Pmt] = payment period, FV = future value, and [Type] = when the payment is made (if nothing is specified then it means that the payment has been made at the end of the period)

Example For a business, its future value of an investment is US\$1000 in 2019. The payment has been made yearly. The interest rate is 10% p.a. Find out the PV (see Fig. 8.57)

Net Present Value (NPV) Net present value is calculated as the sum of all positive and negative cash flows over the years. Its syntax is NPV (Rate, Value 1, [Value 2], [Value 3] ...), where, Rate = discount rate for a period and Value 1, [Value 2], [Value 3] ... = positive or negative cash flows where, negative values would be treated as payments and positive values as inflows

FV = FVSCHEDULE(C4,C5,C7)					
A	B	C	D	E	F
1					
2					
3					
4	Principal	1000			
5	2017	7%			
6	2018	9%			
7	2019	6%			
8					
9	FVSCHEDULE	1236.278			
10					
11					

Figure 8.56 FVSchedule

PV = PV(C2,C3,C4,C5)					
A	B	C	D	E	F
1					
2	RATE	10%			
3	NPER	3			
4	PMT	1			
5	FV	-1000			
6					
7	PV	\$748.83			
8					
9					

Figure 8.57 Present value

NPV = NPV(C2,C3,C7)			
A	B	C	D
1			
2	Rate of Discount	5%	
3	Initial Investment	1000	
4	Return from first year	300	
5	Return from second year	500	
6	Return from third year	600	
7	Return from fourth year	\$700.00	
8			
9	NPV	\$793.74	
10			
11			

Figure 8.58 NPV

Example Consider the data given for company X in Fig. 8.58. Calculate the NPV

XNPV The XNPV function is similar to the NPV function, but with a difference that this function is used when the payment and income are not periodic. Specific dates are mentioned for each payment and income. The syntax for XNPV function is XNPV (Rate, Values, Dates), where, Rate = discount rate for a period, Values = positive or negative cash flows, and Dates = specific dates

Example Consider the data given in Fig. 8.59 and calculate the XNPV.

XNPV = XNPV(C2,C3,C5,C6,D3,D6)				
A	B	C	D	E
1				
2	Rate of Discount	5%		
3	Initial Investment	1000		10/29/2011
4	Return from first year	300		4/1/2012
5	Return from second year	500		2/28/2013
6	Return from third year	600		6/30/2014
7	Return from fourth year	700		9/30/2015
8	Return from fifth year	900		12/31/2016
9				
10	XNPV	\$1506.08		
11				

Figure 8.59 XNPV

PMT PMT means periodical payment required to pay off for a particular period of time with a constant interest rate. The syntax of PMT is $\text{PMT}(\text{Rate}, \text{Nper}, \text{PV}, [\text{FV}], [\text{Type}])$, where, Rate = interest rate period, Nper = number of periods, PV = present value, [FV] = an optional argument which is about the future value of a loan (if nothing is specified then FV is taken as 0), and [Type] = when the payment is made (if nothing is specified then it is considered that the payment has been made at the end of the period).

Example A business needs to pay US\$1000 in full in 3 years. Interest rate is 10% p.a. and the payment needs to be done yearly. Find out the PMT (see Fig. 8.60).

PPMT PPMT is same as PMT with a difference that PPMT calculates payment on principal with a constant interest rate and constant periodic payments. The syntax of PPMT is $\text{PPMT}(\text{Rate}, \text{Per}, \text{Nper}, \text{PV}, [\text{FV}], [\text{Type}])$, where, Rate = interest rate period, Per = the period for which the principal is to be calculated, Nper = number of periods, PV = present value, [FV] = an optional argument which gives the future value of a loan (if nothing is specified then FV is taken as 0), and [Type] = when the payment is made (if nothing is mentioned, it is considered that the payment has been made at the end of the period).

Example A business needs to pay US\$1000 in full in 3 years. Interest rate is 10% p.a. and the payment needs to be done yearly. Find out the PPMT in first year and second year (see Fig. 8.61).

Internal rate of return (IRR) The IRR function is used to determine if a new project or investment is profitable or not. If IRR is more than the acceptable rate, then it is considered to be profitable, else not. The syntax of IRR is $\text{IRR}(\text{Values}, [\text{Guess}])$, where, Values = positive or negative cash flows and [Guess] = an assumption of what you think IRR should be.

Example Use the IRR function on the data given in Fig. 8.62.

	A	B	C	D	E	F
1						
2			RATE	10%		
3			NPER	3		
4			PV	1000		
5						
6			PMT	\$402.11		
7						
8						

Figure 8.60 PMT

	A	B	C	D	E	F
1						
2			RATE	10%		
3			PER	1		
4			NPER	3		
5			PV	1000		
6						
7			PPMT	111		
8						
9						

Figure 8.61 PPMT

	A	B	C	D
1				
2			Initial Investment	-1000
3			Return from 1st year	300
4			Return from 2nd Year	400
5			Return from 3rd Year	500
6			Return from 4th Year	600
7				
8			IRR	25%
9				
10				

Figure 8.62 IRR

Modified internal rate of return (MIRR) Modified internal rate of return signifies that the investment is profitable and is used in business. MIRR is calculated by assuming NPV as zero. The syntax for using MIRR is $\text{MIRR}(\text{Values}, \text{Finance rate}, \text{Reinvestment rate})$, where, Values = positive or negative cash flows, Finance rate = interest rate paid for the money used in cash flows, and Reinvestment rate = interest rate paid for reinvestment of cash flows.

Example Calculate MIRR for the data given in Fig. 8.63.

$f_x = \text{MIRR}(C2:C6, J7, C8)$

	A	B	C	D	E
1					
2		Initial Investment	-1000		
3		Return from 1st Year	300		
4		Return from 2nd Year	400		
5		Return from 3rd Year	500		
6		Return from 4th Year	600		
7		Finance Rate	18%		
8		Reinvestment Rate	12%		
9					
10		MIRR		20%	

Figure 8.63 MIRR

XIRR The XIRR function is used to find the IRR that has specific dates of cash flow. The syntax for the XIRR function is $\text{XIRR}(\text{Values}, \text{Dates}, [\text{Guess}])$, where, Values = positive or negative cash flows, Dates = specific dates, and [Guess] = an assumption of what you think IRR should be.

Example Calculate XIRR on the data given in Fig. 8.64.

$f_x = \text{XIRR}(C2:C6, D2:D6, 0.1)$

	A	B	C	D	E
1					
2		Initial Investment	1000	10/29/2011	
3		Return from 1st year	300	3/31/2012	
4		Return from 2nd year	400	6/30/2013	
5		Return from 3rd year	500	9/30/2014	
6		Return from 4th year	600	2/28/2015	
7					
8		XIRR		31%	
9					
10					

Figure 8.64 XIRR

NPER NPER function is used to specify the number of periods one requires to pay off the loan. Its syntax is $\text{NPER}(\text{Rate}, \text{PMT}, \text{PV}, [\text{FV}], [\text{Type}])$, where, Rate = interest rate per period, PMT = amount paid per period, PV = present value, [FV] = an optional argument which specifies the future value of a loan (if nothing is specified then FV is considered as 0), and [Type] = when the payment is made (if nothing is specified then it means that the payment has been made at the end of the period).

Example A company X pays US\$200 per year for a loan of US\$1000. The interest rate is 10% p.a. and the payment needs to be done yearly. Find out the NPER (see Fig. 8.65).

$f_x = \text{NPER}(C2, C3, C4)$

	A	B	C	D	E	F
1						
2		RATE	10%			
3		PMT	-200			
4		PV	1000			
5						
6		NPER		7.272541		
7						
8						

Figure 8.65 NPER

RATE The RATE function is used to calculate the interest rate needed to pay off the loan in full for a given period of time. Its syntax is $\text{RATE} = (\text{NPER}, \text{PMT}, \text{PV}, [\text{FV}], [\text{Type}], [\text{Guess}])$, where Nper = number of periods, PMT = amount paid per period, PV = present value, [FV] = an optional argument which is about the future value of a loan (if nothing is specified then it is taken as 0), [Type] = when the payment is made (if nothing is specified then it is taken as if the payment has been made at the end of the period), and [Guess] = an assumption of what you think RATE should be.

Example A business pays US\$250 per year for a loan of US\$1000 for 6 years and the payment needs to be done yearly. Calculate the RATE (see Fig. 8.66).

Figure 8.66 RATE

		fx =RATE(C2,C3,C4)		
	A	B	C	D
1				
2		NPER	6	
3		PMT	-250	
4		PV	1000	
5				
6		RATE	13%	
7				
8				

EFFECT The EFFECT function is used to calculate the effective annual interest rate when the nominal interest rate and the number of compounding per year is given. The syntax for this function is $\text{EFFECT} = (\text{Nominal Rate}, \text{NPERY})$, where, Nominal Rate = nominal interest rate and NPERY = number of compounding per year.

Example Calculate the effective annual interest rate for a payment that needs to be paid with a nominal interest rate of 14% with the number of compounding per year being 8 (see Fig. 8.67).

Figure 8.67 EFFECT

		fx =EFFECT(C2,C3)		
	A	B	C	D
1				
2		Nominal Rate	14%	
3		NPERY	8	
4				
5		EAI	0.148882	
6				

NOMINAL The nominal rate for the year is calculated using the effective annual rate and the number of compounding periods per year. The syntax for this function can be given as $\text{NOMINAL} = (\text{Effect Rate}, \text{NPERY})$, where, Effect Rate = effective annual interest rate and NPERY = number of compounding per year.

Example Calculate the nominal rate for a year when a business has to pay with an effective interest rate of 10% when the number of compounding per year is 12 (see Fig. 8.68).

Figure 8.68 NOMINAL

		fx =EFFECT(C2,C3)		
	A	B	C	D
1				
2		Effect Rate	10%	
3		NPERY	12	
4				
5		EAI	0.104713	
6				

SLN The SLN function is used to calculate depreciation via straight line method. The syntax for SLN function can be given as $\text{SLN} = (\text{Cost}, \text{Depreciated value}, \text{Life})$, where, Cost = cost of asset when bought, Depreciated_Value = value of asset after depreciation, and Life = number of periods over which the asset is being depreciated.

Figure 8.69 SLN

		fx =SLN(C2,C3,C4)		
	A	B	C	D
1				
2		COST	10000	
3		DEPRICATED VALUE	2000	
4		LIFE	10	
5				
6		SLN	\$800.00	
7				

Example The initial cost of machinery is US\$10000. After being used for 10 years, its current value is US\$2000. Find depreciation charged per year (see Fig. 8.69).

Data Table A data table is used to perform what-if analysis. In a data table, one or two inputs are varied in some of the cells and the requirement is to come up with different answers to find a solution to a problem. There are two types of data tables: one-variable data tables and two-variable data tables. A one-variable data table is used to see how different values of one variable (in one or more formulae) will change the results of those formulae. In simpler terms, one-variable data table is used to determine how changing one input changes any number of outputs.

Example If there is a loan of ₹10,00,000 for a tenure of 20 years, you can use a data table to know the monthly payments (EMI) for varied interest rates. In this example (Fig. 8.70), we will use the one-variable (the variable here being the rate) data table.

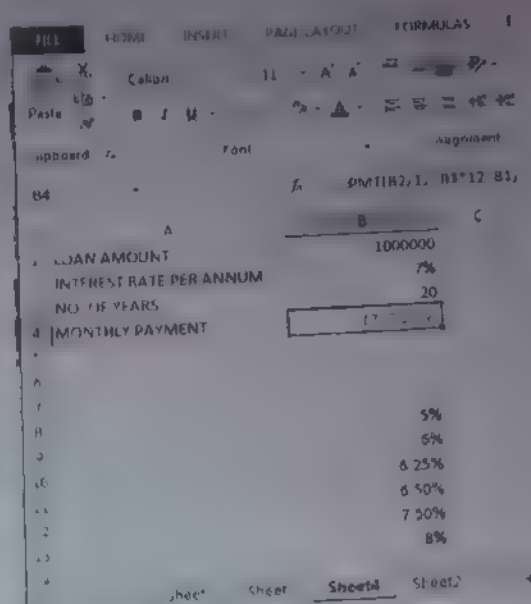


Figure 8.70 One-variable data table

Step 1: List all the required values as shown in the table in Fig. 8.70

Step 2: Use the PMT function to calculate the EMI.

Step 3: Type the list of interest values that may vary as shown in Fig. 8.70

Step 4: Note that there is an empty row above the interest rate values. This row is for the formulae that you want to use. Type PMT function in the cell one row above and one cell to the right of the column of values (Fig. 8.71)

Step 5: Select the range of cells that contains the formulae and values that you want to substitute, that is B7:C13.

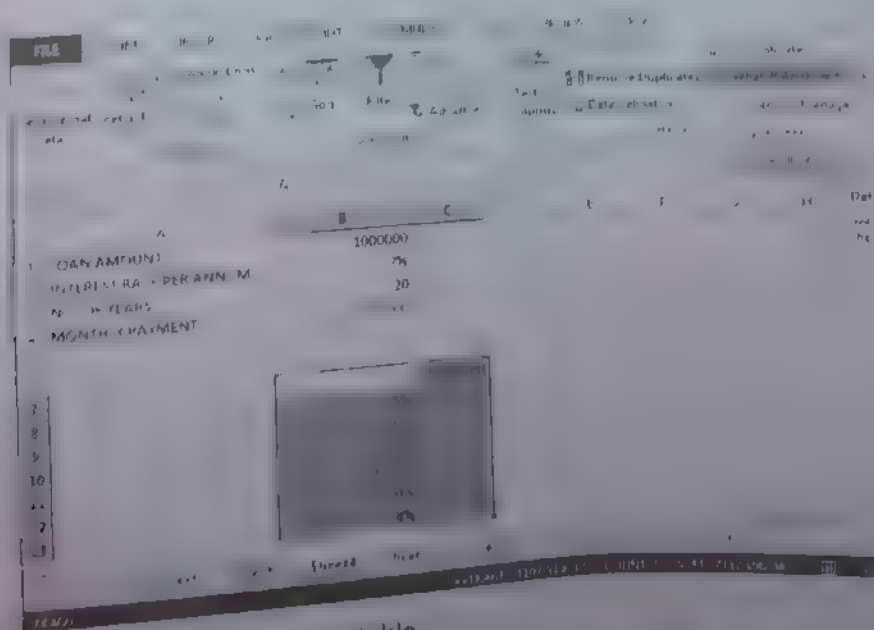


Figure 8.71 Using data table

Step 6: On the DATA tab, click What if Analysis in the Data Tools group and select Data Table from the dropdown list.

Step 7: In the dialog box that appears, click in the Column input cell box.

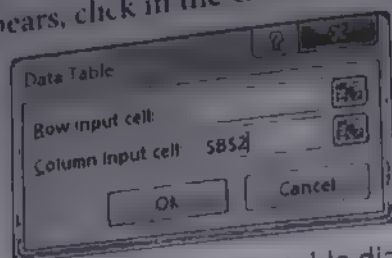


Figure 8.72 Data table dialog box

Step 8: Click the cell B2 (which has the interest rate). The cell reference will be added. Click on OK.

Two-variable data tables A two-variable data table is used to study how different values of two variables in a formula will change the results of that formula. For example, if there is a loan of ₹10,00,000 and you want to know how different combinations of interest rates and loan tenures will affect the monthly payment (EMI). Then you can use a two-variable data table as follows.

Step 1: Perform Steps 1 to 3 as given in case of one-variable data table.

Step 2: In the empty cell, enter the reference of the cell containing the formula (of B4).

Step 3: Select the range of cells that contains the formulae and values that you want to substitute, that is B8:H13.

Step 4: On the DATA tab, click What-if Analysis in the Data Tools group and select Data Table in the dropdown list.

Step 5: In the dialog box that appears click the cells—interest rate (B2) for row as the interest rate will vary in a row and loan amount (B1) for column, as shown in Fig. 8.73.

Step 6: Click on OK.

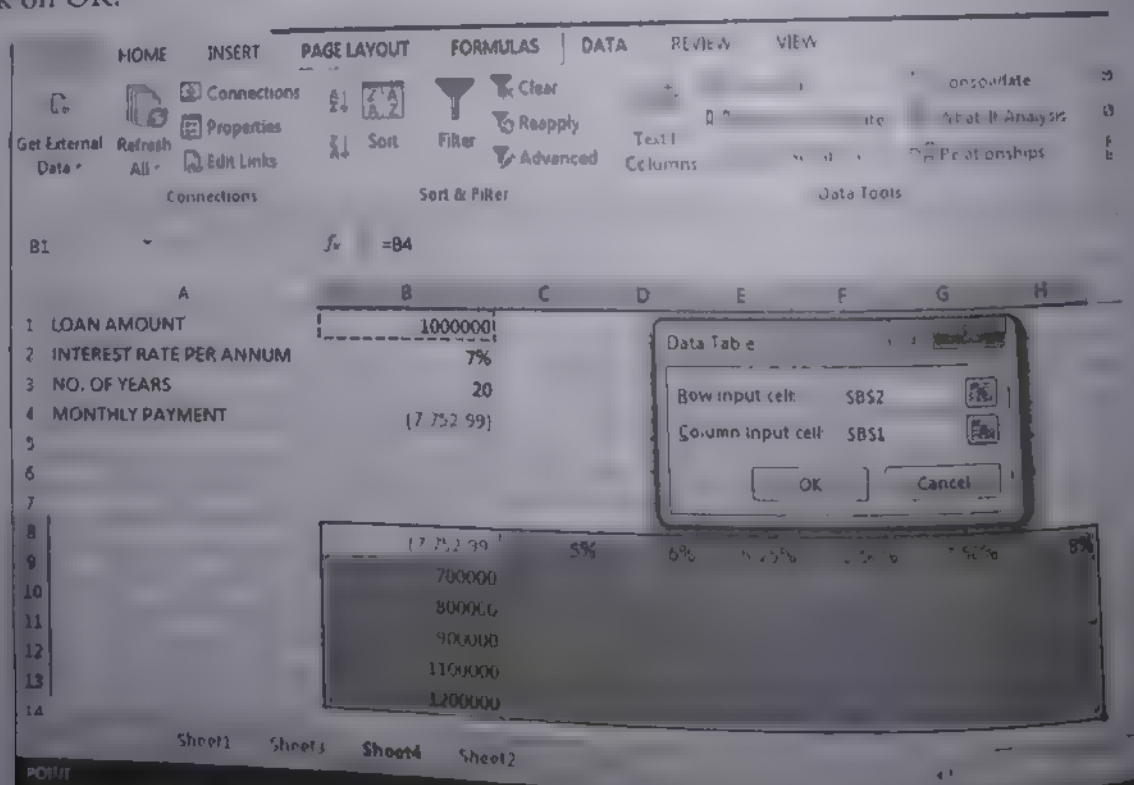


Figure 8.73 Using two-variable data table

Lookup and Reference Functions

The Lookup function is used to search one column of data and find data in the corresponding row. For example, if you are looking for a column with student's Roll No., then LOOKUP() can find student roll no 5678 and then return data from that same row.

Vertical Lookup (VLOOKUP) This function looks for a value in the leftmost column of a table, and then returns a value in the same row from the specified column. To use this function, follow the steps given below

Step 1: Click on a cell.

Step 2: Click on Formulas tab.

Step 3: Click on Insert Function.

Step 4: From the Insert Function dialog box, click on the arrow next to Select a category and select Lookup & Reference as shown in Fig. 8.74

Step 5: Select the VLOOKUP function and fill in the values as specified in Fig. 8.75

The lookup value is one which you are searching (A1), table Array (is given from F5:F9), the value of which column of the table is required (second here) and the last value should be FALSE to return an exact match.

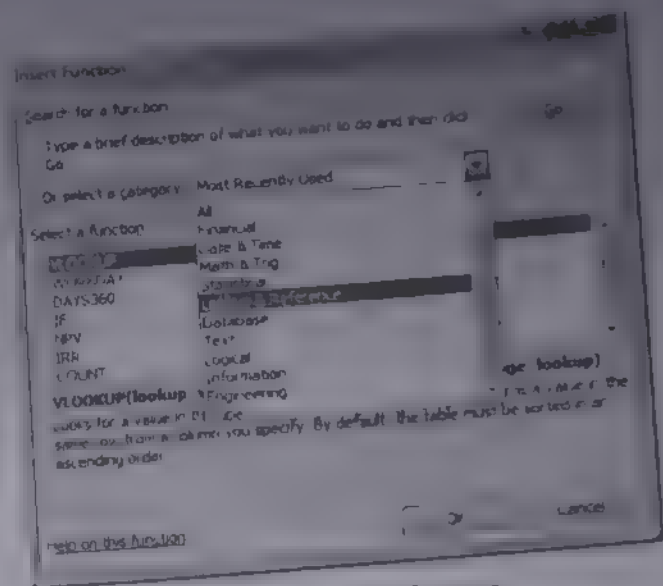


Figure 8.74 Lookup and Reference

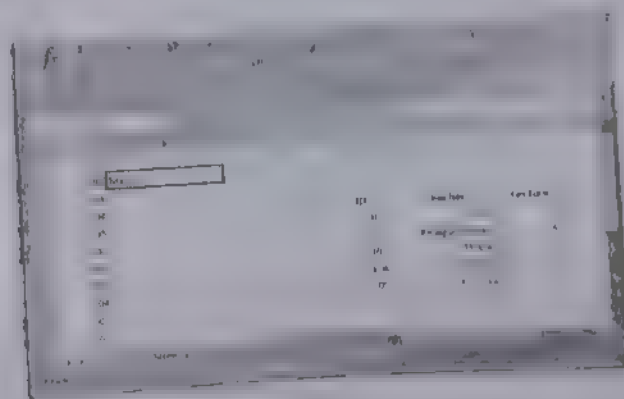


Figure 8.75 VLOOKUP

Horizontal Lookup (HLOOKUP) This function works in the same way as VLOOKUP. The difference is that here the table is specified horizontally rather than vertically as shown in Fig. 8.76.

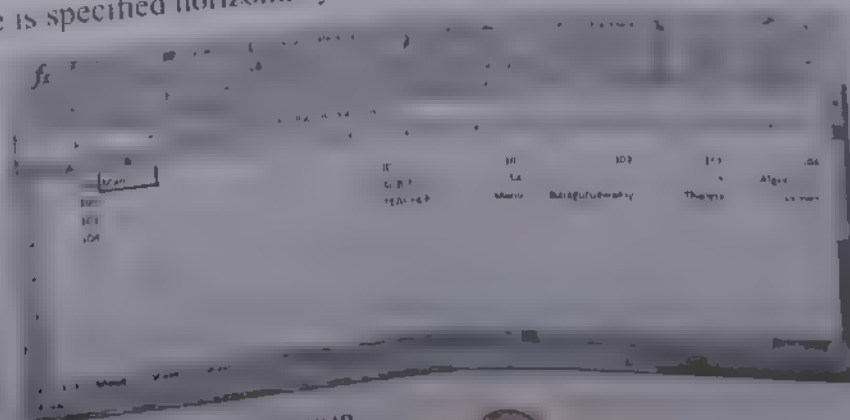


Figure 8.76 HLOOKUP

Match This function returns the position of a value in the specified range. In Fig. 8.77, you can see that we are looking for a value Goransh. The result is written in C1 which is 5 because Goransh is present at the location 5 in array of cells A1:A9. The last argument is optional. The value 0 indicates that an exact match should be found.

Index This function returns a specific value in a range. This range can be either one dimensional or two dimensional. For example, look at Fig. 8.78 which is a two dimensional range (since it has two columns).

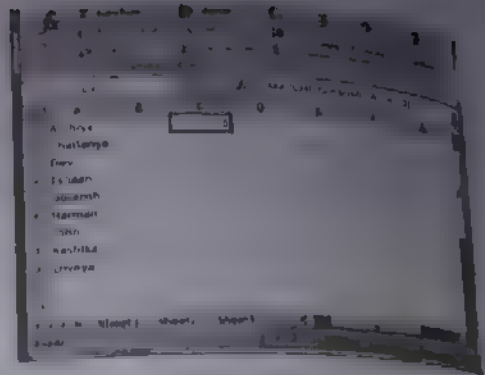


Figure 8.77 MATCH

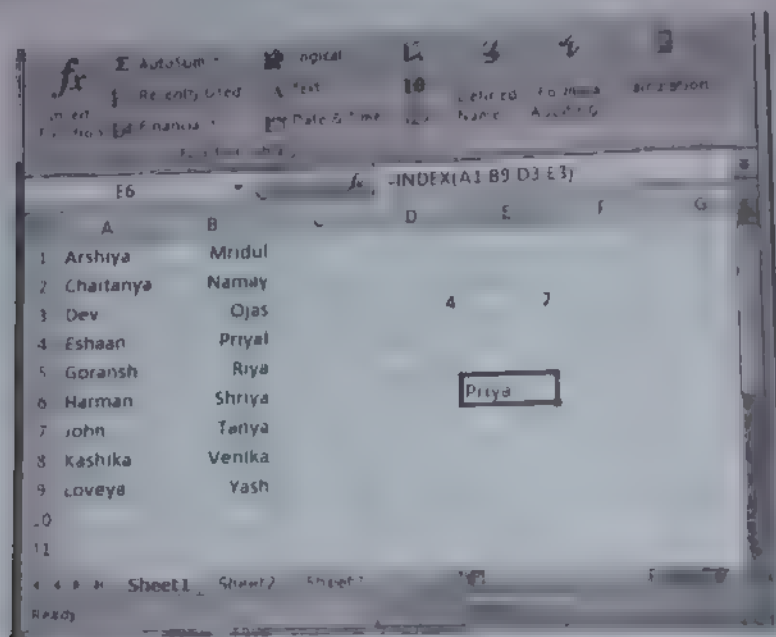


Figure 8.78 INDEX

Now specify the row number of the value in one cell (D3 in one cell) and the column number in another cell (E3, here). Click on any cell (in which you want the value of the cell whose row and column number has been specified) and insert the formula. Click on the Insert tab.

Choose This function gives a value from a list of values based on the specified position. The position is specified as the first argument (refer Fig. 8.79).

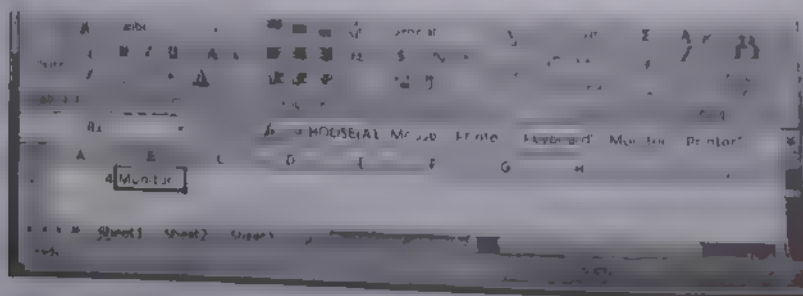


Figure 8.79 CHOOSE

Information Functions Information functions, as the name suggests is used to get information about the data in cell(s). Some frequently used information functions are given in Table 8.3.

Table 8.3 Information functions

Function	Purpose
CELL	Gives information about the formatting, location, or contents of a cell
ERROR TYPE	Gives a number indicating the type of error that has been generated
INFO	Gives information about the current operating environment
ISBLANK	Returns TRUE if the value is blank
ISERR	Returns TRUE if the value is any error value except #N/A
ISNA	Returns TRUE if the value is the #N/A error value
ISERROR	Returns TRUE if the value is any error value
ISEVEN	Returns TRUE if the number is even
ISODD	Returns TRUE if the number is odd
ISLOGICAL	Returns TRUE if the value is a logical value
ISNONTTEXT	Returns TRUE if the value is not text
ISNUMBER	Returns TRUE if the value is a number
ISREF	Returns TRUE if the value is a reference
ISTEXT	Returns TRUE if the value is a text
N	Returns a value converted to a number

Using the Formulas Tab

In Excel, there are two ways to use formulas. First, type the formula. We have already seen this way. Second, use the Formulas tab, the steps given below use the AutoSum option in the Formulas tab.

Step 1: Select the cell.

Step 2: Click on Formulas tab.

Step 3: Click on AutoSum.

Step 4: The range of cells is automatically selected. You can change the selection.

Step 5: Press Enter key to see the answer. To see all in built functions in Excel, click on Formulas tab and select Insert Function. A dialog box will appear from which you can choose a particular function as shown in Fig. 8.80

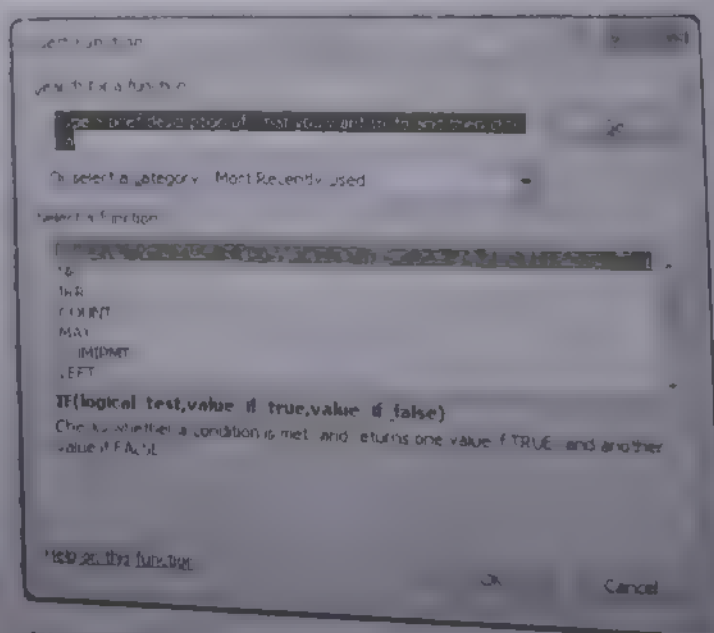


Figure 8.80 Inserting function

Database Functions

Database functions help users to work within a database (i.e., a large number of organized data records), stored in Excel. These functions can perform basic calculations such as sum, average, and count. However, database functions also use criteria arguments that allow calculations to be performed only on a subset of the records in database that meets the criteria. Other records in the database are ignored.

For example, the **DSUM** function in an Employee data list can be used to calculate the sum of salaries of all employees who joined after 1 January, 2018. Similarly, the **DCOUNT** function can be used to compute the number of employees working in the Research department. The database functions begin with **D** and have the following syntax: **Function_name(database, field, criteria)**, where,

- **database** specifies the range containing the list (or database). If you specify a column with a text entry, then an error value or 0 will be returned.
- **field** is the number of the column in the data list. It may also contain the column name.
- **criteria** specifies the address of the cell whose is to be evaluated. It may also specify the range of addresses that would act as the criteria.

Table 8.4 lists the various database functions available in Excel and Figs 8.81(a) and (b) show examples of using some of these functions.

Table 8.4 Database functions

Database Functions	
DAVERAGE	Returns the average of selected database entries
DCOUNT	Counts the cells that contain numbers in a database
DCOUNTA	Counts non-blank cells in a database
DGET	Extracts from a database a single record that matches the specified criteria
DMAX	Returns the maximum value from selected database entries
DMIN	Returns the minimum value from selected database entries
DPRODUCT	Multiplies the values in a particular field of records that match the criteria in a database
DSTDEV	Estimates the standard deviation based on a sample of selected database entries
DSTDEVP	Calculates the standard deviation based on the entire population of selected database entries
DSUM	Adds the numbers in the field column of records in the database that match the criteria
DVAR	Estimates variance based on a sample from selected database entries
DVARP	Calculates variance based on the entire population of selected database entries

A23 f. =DMAX(\$A\$1:\$E\$21,"TOTAL",\$A\$1:\$A\$2)

	A	B	C	D	F	F	G	H
1	ITEM	REGION	UNITS	PRICE	TOTAL			
2	A	NORTH	50	225	11250			
3	B	SOUTH	100	300	30000			
4	C	EAST	75	375	28125			
5	D	WEST	25	290	7250			
6	E	NORTH	30	150	4500			
7	F	SOUTH	45	200	9000			
8	G	EAST	50	275	13750			
9	H	WEST	100	390	39000			
10	I	NORTH	75	165	12375			
11	J	SOUTH	25	225	5625			
12	K	EAST	30	300	9000			
13	L	WEST	45	375	16875			
14	M	NORTH	50	290	14500			
15	N	SOUTH	100	150	15000			
16	O	EAST	75	200	15000			
17	P	WEST	25	275	6875			
18	Q	NORTH	30	390	11700			
19	R	SOUTH	45	165	7425			
20	S	EAST	90	435	39150			
21	T	WEST	110	560	61600			
22								
23								

23 11250

(a)

A23 f. =DSUM(\$A\$1:\$E\$21,5,\$H\$5:\$I\$6)

	A	B	C	D	E	F	G	H
2	A	NORTH	50	225	11250			
3	B	SOUTH	100	300	30000			
4	C	EAST	75	375	28125			
5	D	WEST	25	290	7250			TOTAL
6	E	NORTH	30	150	4500			>10000
7	F	SOUTH	45	200	9000			
8	G	EAST	50	275	13750			
9	H	WEST	100	390	39000			
10	I	NORTH	75	165	12375			
11	J	SOUTH	25	225	5625			
12	K	EAST	30	300	9000			
13	L	WEST	45	375	16875			
14	M	NORTH	50	290	14500			
15	N	SOUTH	100	150	15000			
16	O	EAST	75	200	15000			
17	P	WEST	25	275	6875			
18	Q	NORTH	30	390	11700			
19	R	SOUTH	45	165	7425			
20	S	EAST	90	435	39150			
21	T	WEST	110	560	61600			
22								
23								
24								

23 308325

(b)

Figure 8.81 Using database functions (a) DMAX (b) DSUM

8.5 PIVOT TABLES

Pivot tables is an important feature of Excel which helps you to extract the significance from a large, detailed data set. A pivot table is a special type of summary table that is a great tool for summarizing values in a table without writing any formulas to perform the calculations. You can even play around with the arrangement of the summarized data randomly as and when required by simply rotating the row and column headings. Follow the steps given below to insert a pivot table:

Step 1: Select the data set.

Step 2: Click on the Insert tab.

Step 3: Click on Pivot table command in the Tables group.

Step 4: From the window you may choose the location where you want the table to be inserted. Click on OK as shown in Fig. 8.82.

Step 5: A blank worksheet for the pivot table will be added which displays a PivotTable Field List task pane on the right side of the worksheet area. This pane is divided into two areas: the Choose Fields to Add to Report list box with the names of all the fields in the source data for the pivot table and an area divided into four drop zones (Report Filter, Column Labels, Row Labels, and Values) at the bottom. Drag field name(s) from the Choose Fields to Add to Report list box and dropping it in one or more of the four areas below. You can also select the fields by placing a checkbox next to the field name (refer Fig. 8.83).

Excel allows you to apply Sort & Filter commands on the result set. You can even right click on a cell and using Value Field Settings option choose different types of actions as shown in Fig. 8.84.

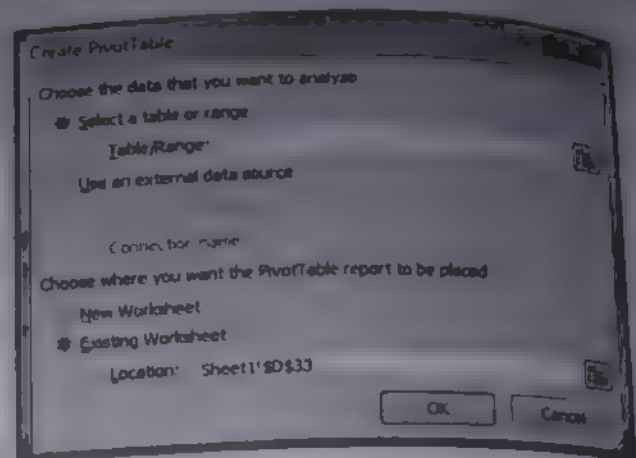


Figure 8.82 Pivot table

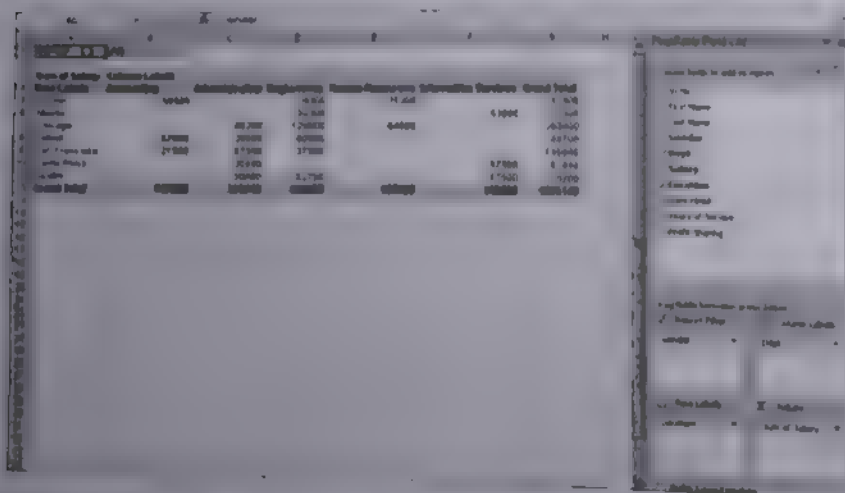


Figure 8.83 Choosing fields

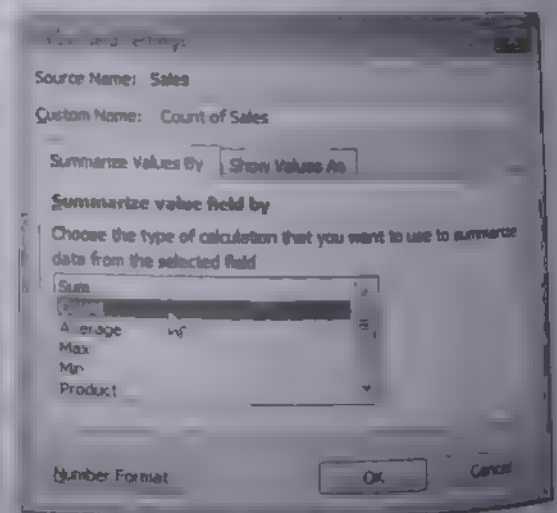


Figure 8.84 Value field settings

8.5.1 Charts

Like formulas and functions, a chart is an integral feature of MS Excel. It is also known as a graph and is used to visually represent numeric values. Displaying data in a chart makes it more understandable

Excel has tools to help you create customized charts. MS Excel is, therefore, a very popular software for summarizing numbers and analysing data. You can draw different types of charts in MS Excel (refer Fig. 8.85). These charts are,

Column A column chart is used to see changes in data over a period of time or to see the comparisons among different items.

Bar A bar graph or a bar chart is used to show comparisons among individual data items.

Pie A pie chart is used to show the size of items that make up a data series, proportional to the sum of the items. It is used to display information about only one data series. Pie chart is especially useful when you want to emphasize a significant element in the data.

Line A line chart is drawn to understand the trends in data at regular intervals.

Area An area chart is used to visualize the magnitude of change in data over time.

XY Scatter An XY (scatter) chart displays the relationships among numbers in several data series. For this, it plots two groups of numeric values as one series of XY coordinates.

Stock It is often used to depict stock price data. However, it can always be used to show other scientific data like temperature changes.

Surface A surface chart is drawn when you want to find optimum combinations between two data sets. For example, in a topographic map, colours and patterns indicate areas that are in the same range of values.

Doughnut A doughnut chart resembles a pie chart as it depicts the relationship of parts to a whole. But unlike the pie chart, a doughnut chart can contain more than one data series.

Bubble In a bubble chart, data is arranged in columns so that X values are listed in the first column and corresponding Y values and bubble size values are listed in adjacent columns.

Radar A radar chart is used to compare the aggregate values of a number of data series.

Terminologies Used in Charts

Chart Area The entire chart area containing the data series, the axes, titles, and legend.

Chart Sheet A separate sheet in the workbook which contains the chart.

Chart Title The title or heading or name of the chart present at the center of the chart area to identify what the data depicts.

Chart Wizard A collection of dialog boxes to simplify the creation of a chart in a step-by-step way to verify the data selection, show possible chart types, and display options for additional items such as titles and legend.

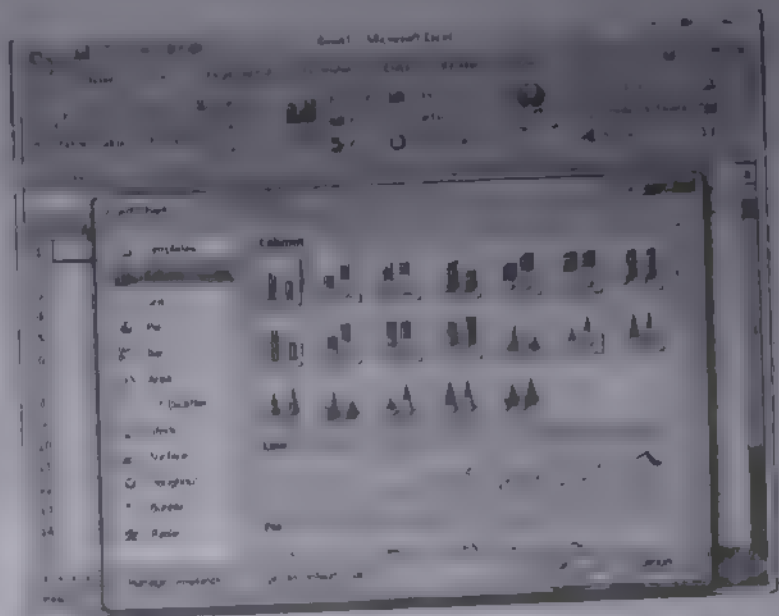


Figure 8.85 Inserting charts

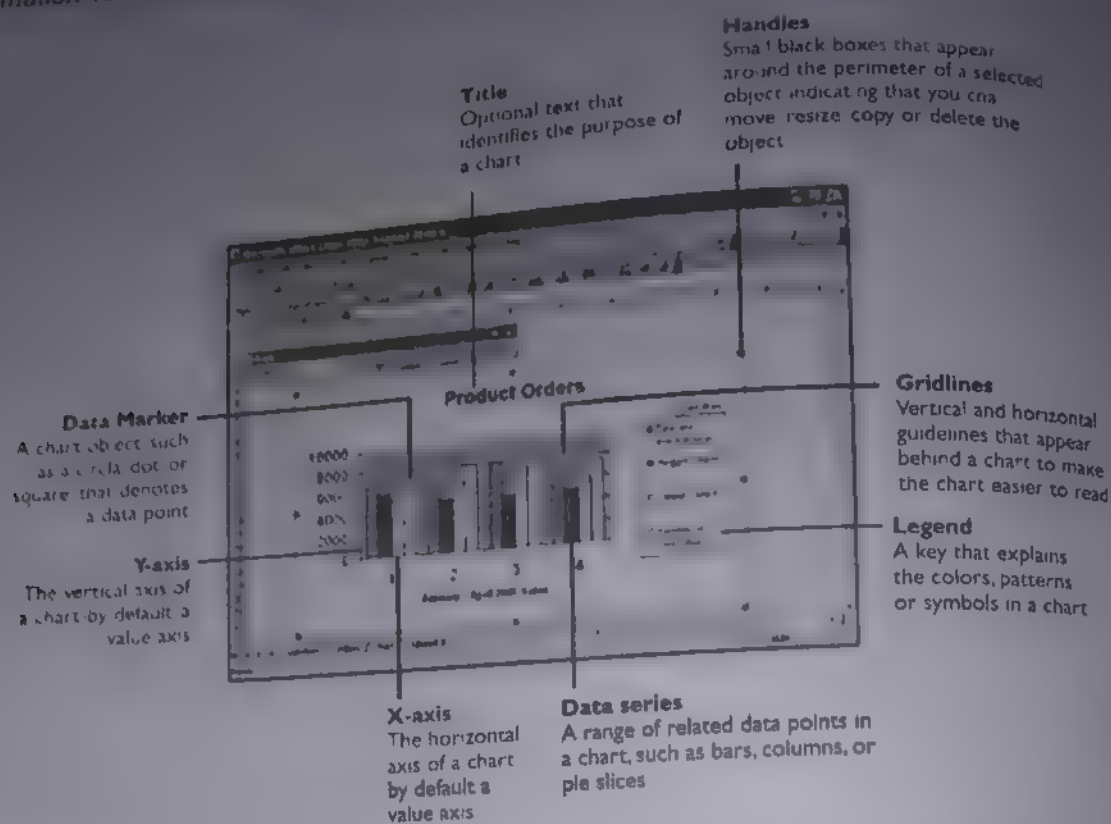


Figure 8.86 Terminologies related to charts

Data Label Value Name assigned to individual data points. Display on chart is optional.

Data Points Values from cells on worksheets that are displayed as bars, lines, columns, pie slices, or other shapes in the chart

Data Series A group of data points. For example, the entire line in a chart.

Gap Width A value that controls the spaces between bars on the chart.

Gridline Lines across plot area. These lines help to easily make reference back to the axes.

Legend Box Identifies data series by names and symbols.

Plot Area The portion of the chart which is bound by the vertical and horizontal axes and their opposite sides.

X-axis Title The title or name to identify what data X-axis is showing.

Y-axis Title Title or name to identify what data the Y-axis is showing.

Printing a Chart

Before printing a chart, the position at which the chart will be printed on the page can be adjusted by moving and sizing the chart. Follow the steps given below to print a chart along with the worksheet data.

Step 1: Click the worksheet that contains the chart that you want to print.

Step 2: Click on Page Layout or Page Break Preview in the Workbook Views group in the View tab

Step 3: If you want to move the chart, just drag it to the desired position on the page that you want to print.

Step 4: To resize the chart, either click on it and drag the sizing handles or enter the specific size in the Shape Height and Width box in the Size group in the Format tab.

Step 5: Click on the File menu and then on Print.



To see how the chart will look on the printed page, you can first click on the Preview command to see a preview of the chart.

To print a chart without the worksheet data, simply click on the chart and then on Print command in the File menu.

Lab Activity: Creating Column Chart

Step 1: Select the data for which you want to create chart.

Step 2: Click on Insert tab.

Step 3: Click on Charts group and click on the desired type of chart as shown in Fig. 8.87

Step 4: Click on OK to generate the chart

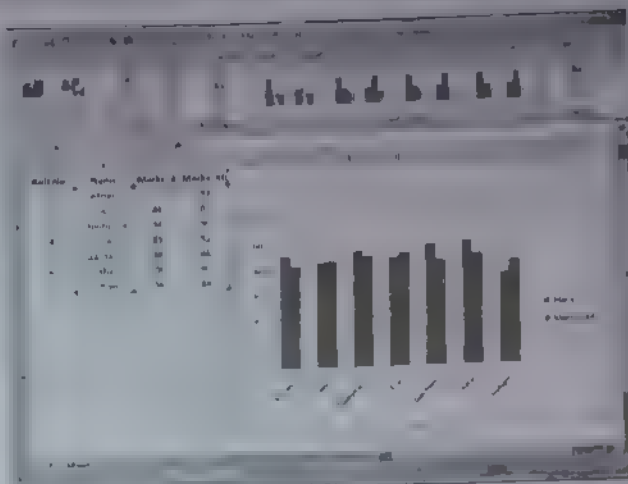


Figure 8.87 Generating charts



Once the chart is created, you can edit it by right-clicking on it.

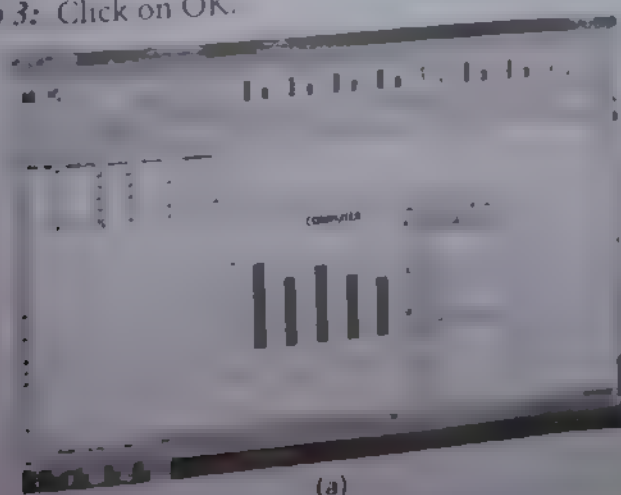
Changing The Data Source

After creating a chart in Excel 2013, you may need to change the data for which the chart has to be drawn. For example, you may draw a chart for Computer Science subject but then later on want to see the graph of Physics. In such a situation, you will have to use the change data source feature. Follow the steps given below to perform this task.

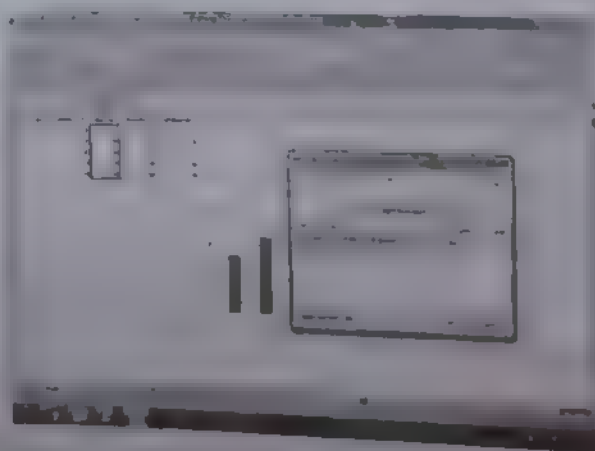
Step 1: Right-click on the chart and click on Select Data as shown in Fig. 8.88a.

Step 2: Click and drag the new data as shown in Fig. 8.88b

Step 3: Click on OK.



(a)



(b)

Figure 8.88 Changing the data source

Inserting Chart Title, Axis Title, and Data Labels in Charts

To further format your chart you can add a title to your chart and its axes, and label the data to make the chart self-explanatory. Follow the steps given below to perform these tasks:

Step 1: Select the chart.

Step 2: Click on the Layout tab.

Step 3: In the Labels group, click on Chart Title, select the particular option and type the title.

Step 4: In the Labels group, click on Axis Title. Select Primary Horizontal Axis Title and type the title. Then select Vertical Axis title and type the vertical title as shown in Fig. 8.89.

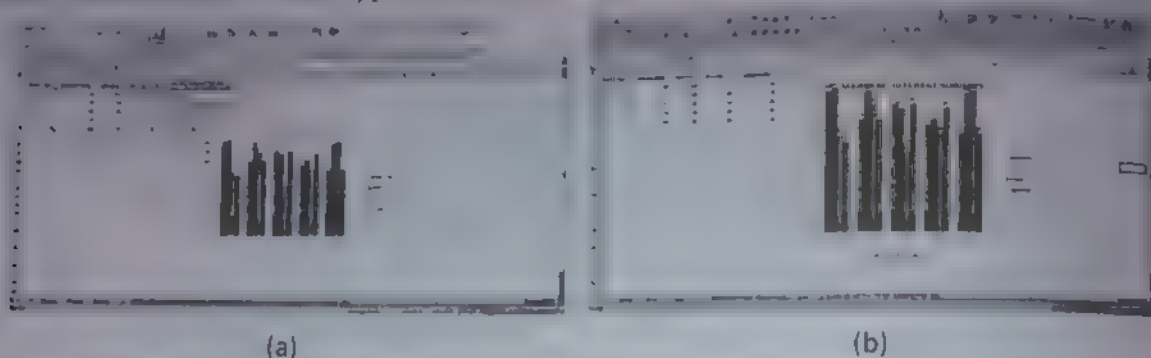


Figure 8.89 Inserting chart title, axis title, and data labels in charts

Step 5: In the Labels group, click on Legend and select an appropriate option.

Step 6: In the Labels group, click on Data Labels and select the Show option.

You can even add the data table in the chart by clicking on the Data Table button in the Labels group.

8.6 OTHER APPLICATIONS OF SPREADSHEETS

Spreadsheets are immensely useful in a host of other areas such as ratio analysis, payroll statements, capital budgeting, and many more. A few of these applications are discussed in this section.

8.6.1 Ratio Analysis

Ratio analysis is a technique to compare the different aspects of a business's finances. It is specifically done to test a business's efficiency, liquidity, profitability, and solvency. Ratio analysis is even more important when it is used to compare similar companies in the same industry.

Current position analysis uses liquidity ratios such as working capital, current ratio, and acid-test ratio. Liquidity ratios measure the ability of a company to meet its currently maturing debts determined from the balance sheet. These ratios are especially of interest to short-term creditors. *Working capital* is measured as follows:

$$\text{Working Capital} = \text{Current assets} - \text{Current liabilities}$$

That is, it is a measure that specifies how easily a company can meet its current obligations. Although working capital is a good measurement for making intra-period comparisons, it is not of much use while comparing companies of different sizes or industries.

Current ratio, also known as *working capital ratio*, compares a company's current assets to its current liabilities. The general rule is that the current ratio should be at least 2:1. This ratio actually indicates a company's solvency to a lending corporation than a company with a large working capital.

$$\text{Current Ratio} = \text{Current assets} / \text{Current liabilities}$$

Acid test ratio, also known as the *quick ratio*, is a stronger measure of liquidity as it only uses cash and those items that can be quickly converted to cash from the current assets. It is different from the current ratio as it includes items that would be difficult to convert to cash

$$\text{Acid test ratio} = \frac{\text{Cash} + \text{Accounts receivable} + \text{Short-term investments}}{\text{Current liabilities}}$$

Average collection period specifies the number of days it takes a business to convert its account receivables into cash. That is, it gives the average number of days between the date a credit sales is made and the day it is collected

$$\text{Number of Days} = \text{Accounts receivable} \times (\text{Net sales} / 365)$$

Average collection period is important as a company can compare it with the previous years to determine if customers are taking longer to pay back on their accounts or if they are paying them back quicker. No doubt, smaller the number of days, the better it is

Debt ratio compares a company's total debt to its total assets

$$\text{Debt ratio} = \text{Total debt} / \text{Total assets}$$

Some of the other important ratios include the following

$$\text{Debt to equity ratio} = \text{Total liabilities} / \text{Stockholders' equity}$$

$$\text{Price to earnings} = \text{Market price of common stock} / \text{Earnings per share}$$

$$\text{Gross profit margin} = \text{Gross profit} / \text{Net sales}$$

$$\text{Net profit margin} = \text{Net income} / \text{Net stock}$$

$$\text{Return on investment} = \text{Net earnings} / \text{Total assets}$$

$$\text{Return on equity} = \text{Net earnings} / \text{Stockholders' equity}$$

The return on equity ratio indicates how well the company has done with the amount the shareholders have invested

8.6.2 Payroll Statements

A payroll is a company's list of its employees. Correspondingly, the term "payroll statement" of a company details the total amount of money that a company pays to its employees and a record of its employees' salaries, bonus, and taxes

Payroll statements are very crucial for every company small or big because it considerably affects its net income. Even for employees, payrolls are important to ensure that they are paid accurately and timely with the correct withholdings and deductions

8.6.3 Capital Budgeting

Capital budgeting is the process in which a business evaluates and ranks its potential expenses or investments that are significant in amount. These expenditures and investments include projects such as purchasing a new equipment, building a new plant, purchasing delivery vehicles, or investing in a long-term venture. In this process, the project's lifetime cash inflows and outflows are assessed to determine whether the potential returns generated meet its target profit by a certain period. The period is calculated based on the time value of money, the number of years it takes for a project's cash flow to pay back the initial cash investment, an assessment of risk, and other factors

8.6.4 Depreciation



8.6.5 Frequency Distribution and Statistical Parameters

	A	F	F	G	H
1	ROLL NO	MARKS			
2	1	89			
3	2	92			
4	3	78			
5	4	84			
6	5	89			
7	6	92			
8	7	78			
9	8	64			
10	9	77			
11	10	92			
12	11	78			
13	12	64			
14	13	98			
15					
16					
17	NO. OF CHILDREN GETTING 60-70 MARKS			3	
18	NO. OF CHILDREN GETTING 70-80 MARKS			4	
19	NO. OF CHILDREN GETTING 80-90 MARKS			2	
20	NO. OF CHILDREN GETTING 90-100 MARKS			4	

Summary

- Microsoft Excel is a commercial spreadsheet application developed by Microsoft.
- In Excel, a cell is addressed using a combination of its row and column numbers.

- Saving a workbook automatically saves all its worksheets
- A Microsoft Excel cell can hold data of different types like numbers, currency, dates, etc.
- You can change the width of column(s) by dragging the right boundary of the column heading until the column has the desired width.
- The result of formula is displayed in the cell that has been currently selected and contains the formula
- The name of every function is followed by brackets. In between the brackets, the arguments (range of cells) are specified.
- When the data of cell(s) is linked to other cells, then the data is dynamically pulled from the source cell(s) into the linked cell(s).
- The correlation coefficient is a value between -1 and $+1$, which tells us how strongly two variables are related to each other.
- Regression analysis is done for estimating the relationship between a dependent variable and one or more independent variables.
- The VARA function calculates the sample variance of a supplied set of values.
- The FV function is used to find out the future value of a particular investment that has a constant interest rate and periodic payments.
- The FVSCHEDULE function is used to calculate the future value with the variable interest rate.
- The Present Value function is used to find out PV.
- Net Present Value is calculated as the sum of all positive and negative cash flows over the years.
- The XNPV function is similar to the NPV function but with a difference that this function is used when the payment and income are not periodic.
- PMI means periodical payment required to pay off for a particular period of time with a constant interest rate.
- The IRR function is used to determine if a new project or investment is profitable or not. If IRR is more than the acceptable rate, then it is considered to be profitable else not.
- NPER function is used to specify the number of periods one requires to pay off the loan.
- The RATE function is used to calculate the interest rate needed to be paid to pay off the loan in full for a given period of time.
- The EFFECT function is used to calculate the effective annual interest rate when the nominal interest rate and the number of compounding per year is given.
- The SLN function is used to calculate depreciation via straight line method.
- The STDEV function calculates standard deviation of a supplied set of values.
- Data table is used to perform what-if analysis. In a data table, one or two inputs are varied in some of the cells and the requirement is to come up with different answers to find a solution to a problem.
- Database Functions help users to work within a database (i.e. a large number of organized data records), stored in Excel.
- A frequency distribution table in Excel gives a snapshot of how data is spread out.
- Ratio analysis is specifically done to test a business's efficiency, liquidity, profitability, and solvency.
- Payroll statement of a company details the total amount of money that a company pays to its employees and is a record of its employees' salaries, bonus, and taxes.

Glossary

Cell Intersection of a row and a column

Formula An expression that performs a calculation

Function Predefined formulas which are already available in Excel

Multiple-choice Questions

1. Excel is developed by _____.
 (a) Adobe (c) Firefox
 (b) Google (d) Microsoft
2. Excel is a _____ software.
 (a) application (c) graphics
 (b) word processing (d) presentation
3. A file in Excel is saved using the _____ extension.
 (a) pptx (c) xlsx
 (b) docx (d) .shtx
4. In which layout of Excel, pages are displayed as they will appear when printed?
 (a) Normal layout
 (b) Page layout
 (c) Page break view layout
 (d) None of these
5. You can have a maximum _____ columns in your Excel sheet.
 (a) 16,384 (c) 1,048,576
 (b) 655,356 (d) 12,345
6. To print a page, you can press _____ keys
 (a) Ctrl + T (c) Alt + P
 (b) Ctrl + P (d) Alt + T
7. When you open a workbook, Excel gives you _____ worksheets
 (a) 1 (c) 3
 (b) 2 (d) 4
8. _____ is an individual addressable location in an Excel sheet.
 (a) Cell (c) Column
 (b) Row (d) All of these
9. A cell reference or address can also be specified as _____.
 (a) SheetName.Cell
 (b) SheetName&Cell
 (c) SheetName#Cell
 (d) SheetName!Cell
10. Pressing Shift + F11 will _____ a worksheet.
 (a) insert (c) edit
 (b) delete (d) save
11. A column can have a maximum of _____ characters.
 (a) 128 (c) 512
 (b) 255 (d) 1,024
12. The default column width is _____ characters.
 (a) 255 (c) 8.43
 (b) 256 (d) 8.34
13. The maximum row height can be _____ points
 (a) 255 (c) 409
 (b) 512 (d) 72
14. The default row height is _____ points
 (a) 72 (c) 8.43
 (b) 409 (d) 12.75
15. Formulas always begin with a _____ sign.
 (a) ! (c) >
 (b) = (d) -
16. Formulas are displayed in the _____ bar
 (a) Status (c) Title
 (b) Menu (d) Formula
17. Which function is used to find the most frequently occurring number?
 (a) Mean (c) Mode
 (b) Median (d) Count
18. Which function is used to get the current date and time?
 (a) Date (c) Year
 (b) Time (d) Now
19. Round (123.789065, 3) will give the result _____.
 (a) 123 (c) 23.7
 (b) 123.789 (d) 789
20. Excel allows you to nest up to _____ IF functions
 (a) 6 (c) 8
 (b) 7 (d) 9
21. Which function returns the position of a value in the specified range?
 (a) Index (c) Match
 (b) Pos (d) Loc
22. Which chart is used to show comparisons among individual data items?

- (a) Column (c) Pie
(b) Bar (d) Line
23. Which chart is used to compare the aggregate values of a number of data series?
(a) Bubble (c) Surface
(b) Radar (d) Line
24. To create links between data you should use _____ command.
(a) Copy
(b) Paste
(c) Paste special
(d) Paste link
25. Which option is used to change the orientation of the pasted entries?
(a) Transpose (c) Formats
(b) Paste link (d) Values
26. _____ are defined to perform the same task or action on same type of data
(a) Formulas (c) Macros
(b) Functions (d) Charts
27. _____ allow users to record different tasks and apply them over on some other portion of the data
(a) Formulas (c) Macros
(b) Functions (d) Charts
28. _____ provides optional commands and features for Microsoft Excel that is by default not available in Excel
(a) Formulas (c) Macros
(b) Functions (d) Add-ins
29. Which function checks how strongly two variables are related to each other?
(a) Correl (c) VAR
(b) STDEV (d) VARA
30. A correlation coefficient of _____ indicates that if one variable increases, the other also increases
(a) 1 (c) +1
(b) 0 (d) 100
31. _____ is done to estimate the relationship between a dependent variable and one or more independent variables
(a) Correlation analysis
(b) Regression analysis
(c) Standard deviation
(d) Mean and mode analysis
32. The minimum and maximum arguments in VARA function are _____ respectively.
(a) 0 and 255 (c) 2 and 225
(b) 1 and 255 (d) 3 and 255
33. _____ function is used to calculate the future value with variable interest rate.
(a) FV (c) NPV
(b) FVSCHEДУLE (d) PV
34. _____ is calculated as the sum of all positive and negative cash flows over the years.
(a) FV (c) NPV
(b) FVSCHEДУLE (d) PV
35. The _____ function is used to calculate as the sum of all positive and negative cash flows when the payment and income are not periodic
(a) XNPV (c) NPV
(b) FVSCHEДУLE (d) PV
36. The _____ function is used to calculate periodical payment required to pay off for a particular period of time with a constant interest rate
(a) FV (c) NPV
(b) PMT (d) PV
37. The _____ function is used to determine if a new project or investment is profitable or not.
(a) FV (c) NPV
(b) PMT (d) IRR
38. The _____ function is used to find out IRR which has specific dates of cash flow
(a) XIRR (c) NPV
(b) PMT (d) MIRR
39. The _____ function is used to calculate the effective annual interest rate when the nominal interest rate and the number of compounding per year is given.
(a) XIRR (c) EFFECT
(b) RATE (d) MIRR
40. The _____ function is used to calculate depreciation via straight line method
(a) DDEP (c) SLN
(b) NPFR (d) DDEP

41. To see total value on a filtered data, the function is used
 (a) TOTAL (c) SUM
 (b) SUBTOTAL (d) AUTOSUM
42. If only one array is supplied, SUMPRODUCT will simply _____ the items in the array
 (a) sum (c) multiply
 (b) subtract (d) divide
43. All values in the array must be _____
 (a) numeric (c) dates
 (b) text (d) formulae
44. The _____ function is used when you want to include logical values and text representations of numbers in a reference as part of the calculation.
 (a) STDEV (c) STDFVP
 (b) STDEVA (d) STDEV D
45. _____ is used to determine how changing one input changes any number of outputs.
 (a) Regression analysis
 (b) Correlation
 (c) Data table
 (d) STDEV
46. _____ is a table in Excel that gives a snapshot of how data is spread out
 (a) Data table
 (b) Frequency distribution table
 (c) Statistical table
 (d) None of these
47. _____ is measured as current assets - current liabilities
 (a) Working capital
 (b) Current ratio
 (c) Acid-test ratio
 (d) Liquidity ratio
48. _____ compares a company's current assets to its current liabilities.
 (a) Working capital
 (b) Current ratio
 (c) Acid-test ratio
 (d) Liquidity ratio
49. _____ is a stronger measure of liquidity, as it only uses cash and those items that can be quickly converted to cash from the current assets.
 (a) Working capital (c) Acid-test ratio
 (b) Current ratio (d) Liquidity ratio

Answers to Multiple-choice Questions

- 1 (d) 2. (a) 3 (c) 4. (b) 5. (a) 6 (b) 7. (c) 8. (a) 9. (d) 10 (a)
 11 (b) 12 (c) 13 (c) 14. (d) 15 (b) 16 (d) 17. (c) 18. (d) 19. (b) 20 (b)
 21 (c) 22. (b) 23 (b) 24 (d) 25. (a) 26 (c) 27. (c) 28. (d) 29. (a) 30. (b)
 31 (b) 32 (c) 33 (b) 34 (c) 35. (a) 36. (b) 37 (d) 38. (a) 39 (a) 40 (c)
 41 (b) 42 (a) 43 (a) 44 (b) 45 (c) 46 (b) 47 (a) 48 (b) 49 (c)

MS Access: A Database Management System

Syllabus Mapping

Unit

Module II
Unit 9

Creation of tables, multiple table handling, defining relationship [foreign key], simple and conditional queries, types of queries [update, delete, append], forms, reports, introduction to SQL through basic commands.
Applying DBMS in the areas of accounting, inventory, HRM and its accounting, managing the data records of employees, suppliers, and customers.

9.1 INTRODUCTION

Microsoft Access is a database management system (DBMS) from Microsoft that combines the relational Microsoft Jet Database Engine with a graphical user interface and software development tools. It is a member of the Microsoft Office suite of applications, included in the professional and higher editions.

Similar to relational databases, Microsoft Access allows you to link related information easily. It can also import or link directly to data stored in other applications and databases.

Databases in Access are composed of many objects, but the following are the major objects.

- Tables
- Queries
- Forms
- Reports

9.2 CREATING DATABASES

In MS Access, you can create a desktop database either by using a template or create a blank database from scratch.

An Access desktop database is a fully functional relational database management system (RDBMS).

On first opening the MS Access 2013, you will see a list of templates available to create a database from, as shown in Fig. 9.1.

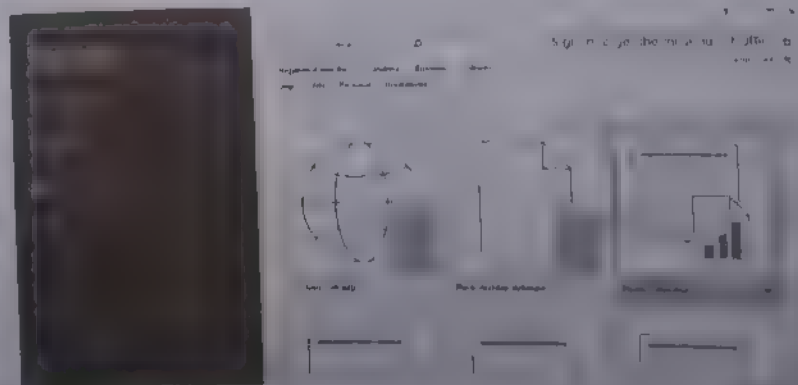


Figure 9.1 Creating a database

9.2.1 Creating Database using Design Template

The simplest way to create a table is to use templates. A template is a ready-to-use database in which tables, queries, reports, and forms are already designed.

Step 1: In Access start screen, click on any template, say Product Inventory in Suggested searches (see Fig. 9.2).

Step 2: A description of the template is displayed and asks you to enter the name of database under File Name.

Step 3: Enter the file name and click on Create.

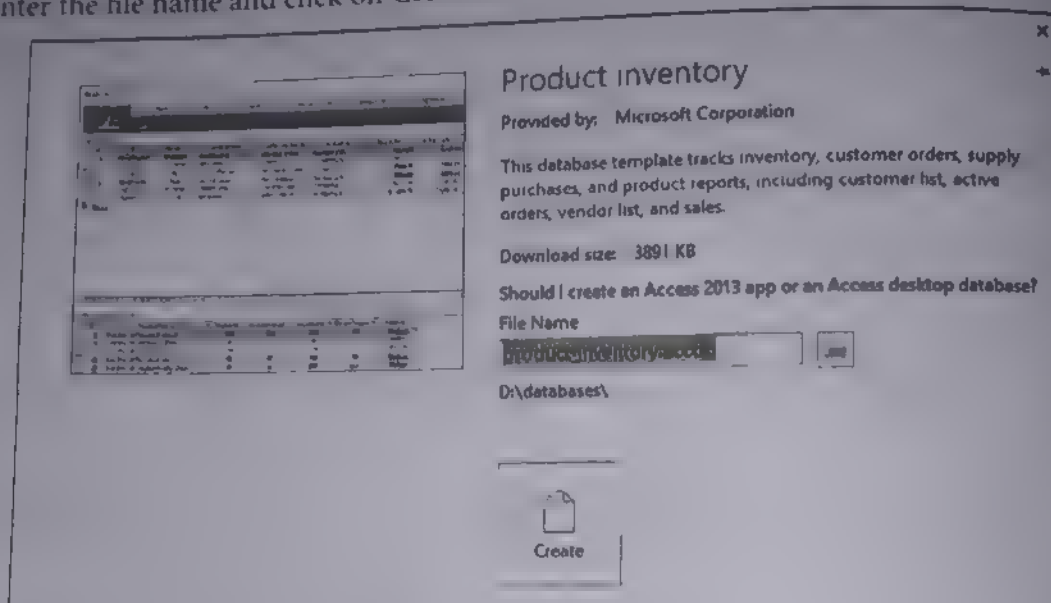


Figure 9.2 Using templates to create a database

Access 2013 creates the database as shown in Fig. 9.3. The navigation pane contains all its objects.

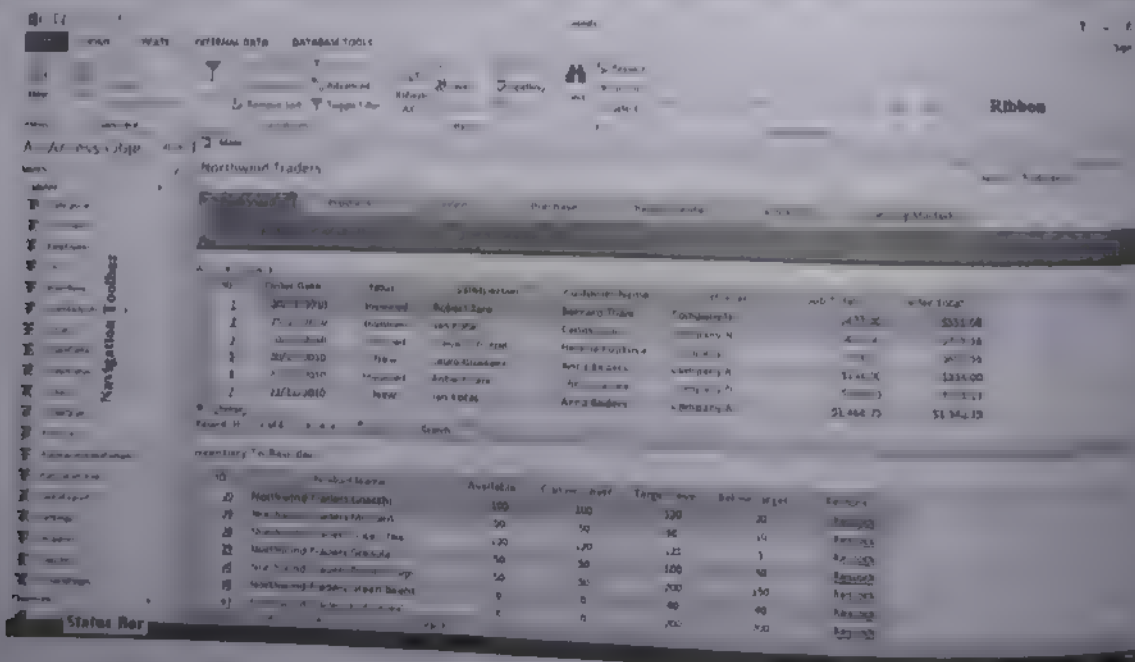


Figure 9.3 Database created using a template

9.2.2 Creating a Blank Desktop Database

Sometimes database requirements can be so specific that using and modifying the existing templates requires more work than just creating a database from scratch. In such a case, we make use of blank database.

Step 1: Open MS Access.

Step 2: Select Blank desktop database. Enter the name and click the Create button (see Fig. 9.4).

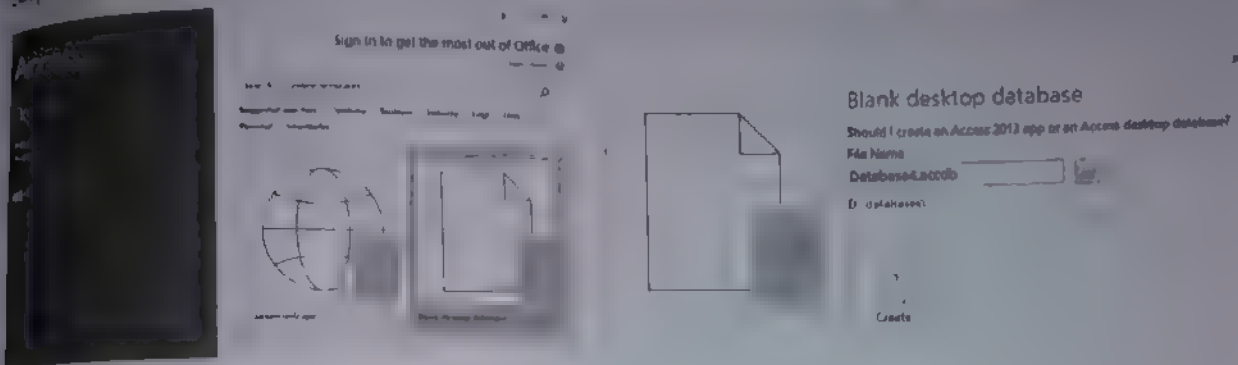


Figure 9.4 Creating a blank desktop database

Step 3: Access will create a new blank database and will open up the table which is also completely blank as shown in Fig. 9.5

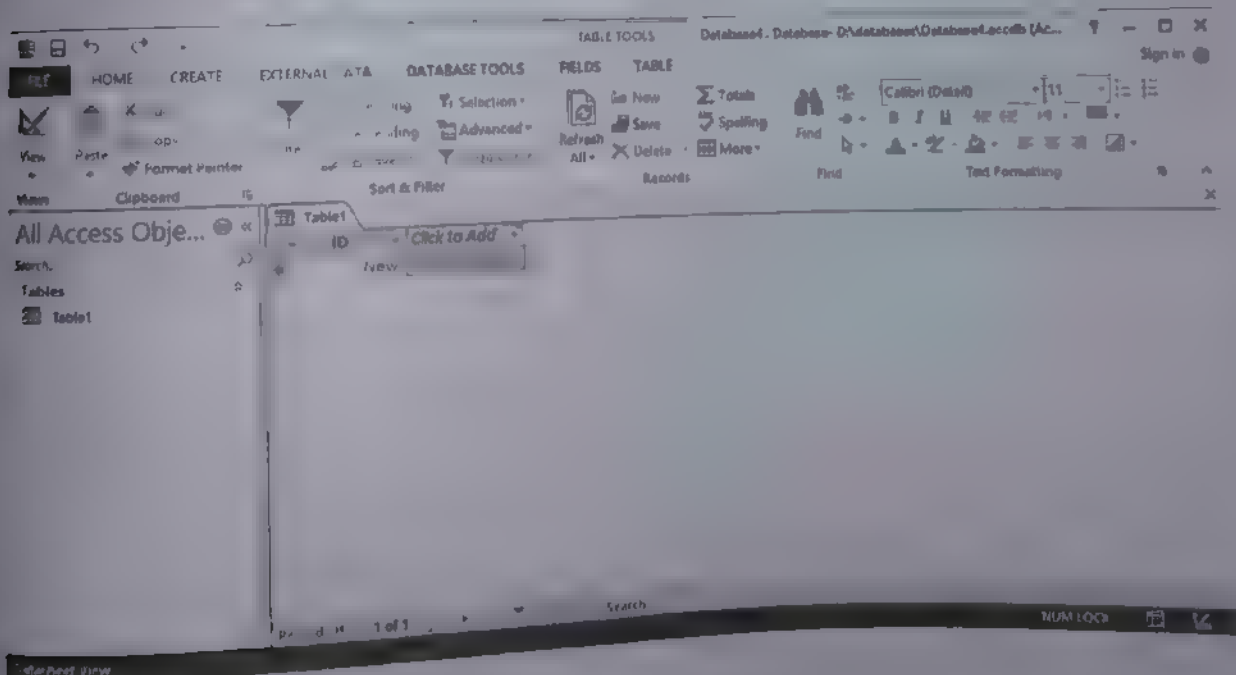


Figure 9.5 A blank database



An Access desktop database has accdb extension. A database can have multiple tables.

9.3 DATA TYPES IN MS ACCESS

A field can store different types of data. Therefore, while adding a field, you must specify its data type. The data type of a field determines the type of data that a field can store. The different data types available in MS Access are listed in Table 9.1.

Table 9.1 Data types in MS Access

Short Text	Stores text. Can also store a combination of text and numbers (like address). The default size is 50 but it can store up to 255 characters.
Text	Can store long text. For example, to give notes or descriptions. It can store 65,535 characters.
Number	Stores numeric data (integers or decimal numbers) that can be used in calculations.
Currency	Stores currencies and displays them in different formats.
Date/Time	Stores date and time and displays them in different formats.
Yes/No	Has only two values—true or false. For example, Adult, Married.
Attachment	This is used to attach external files (such as documents, charts, images, etc.) to a database.
Hyperlink	This is an address link to a folder on a hard disk in your computer or a website on the Internet.
OLE Object	It is used to include pictures, graphs, worksheets, or any object created in another software.
Auto Number	Stores an integer whose value increases or decreases automatically as new records are added or deleted. If the table does not have a primary key, auto number uniquely identifies records.
Calculated	Stores the results of a calculation done on other fields in the same table.
Lookup & Relationship	Displays a list of values from which you can choose a particular value. For example, if we have a field named Marital Status in a table, then to avoid typing text we can simply choose a value from a list of values. This is very helpful when you have a limited set of answers.

The image shows a screenshot of a Microsoft Access form. At the top, there is a label 'Marital Status' followed by a small downward arrow icon. Below this label is a rectangular box representing a dropdown menu. The menu is currently open, displaying three options: 'Married', 'Single', and 'Separated'. Each option is on a new line within the menu box.



An error message will be displayed if you specify the data of one type and store a value of another type.

9.4 TABLES IN MS ACCESS

Table is an object that is used to define and store data. When you create a new table, Access asks you to define fields, also known as column headings. Each field must have a unique name and data type.

Tables contain fields or columns that store different kinds of data, such as a name or an address, and records or rows that collect all the information about a particular instance of the subject, such as all the information about a customer or employee.

9.4.1 Creating Tables in Datasheet View

When you create your database, a blank new table is opened by default in Datasheet View.

Example 9.1 Create a database, say XYZ organization and create a table called *Employee* to store the details of all employees in this organization.

Solution: Step 1: Create a blank database and name it XYZ Organization (Fig. 9.6). Click on create. It will open a blank table in Datasheet View.

The first field in the new table is by default ID, which is the primary key. A primary key uniquely identifies each record in the table. A table can have only one primary key, but this primary key can be made by one or more columns.

Step 2: Select the empty cell below. Click to Add. Enter the required value and press Enter.

Step 3: Enter the data in the remaining fields similarly (see Fig. 9.7).

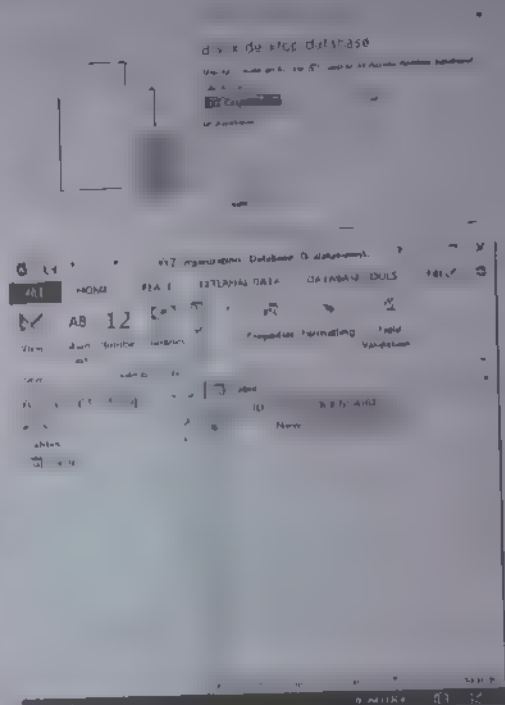


Figure 9.6 Creating database XYZ Organization

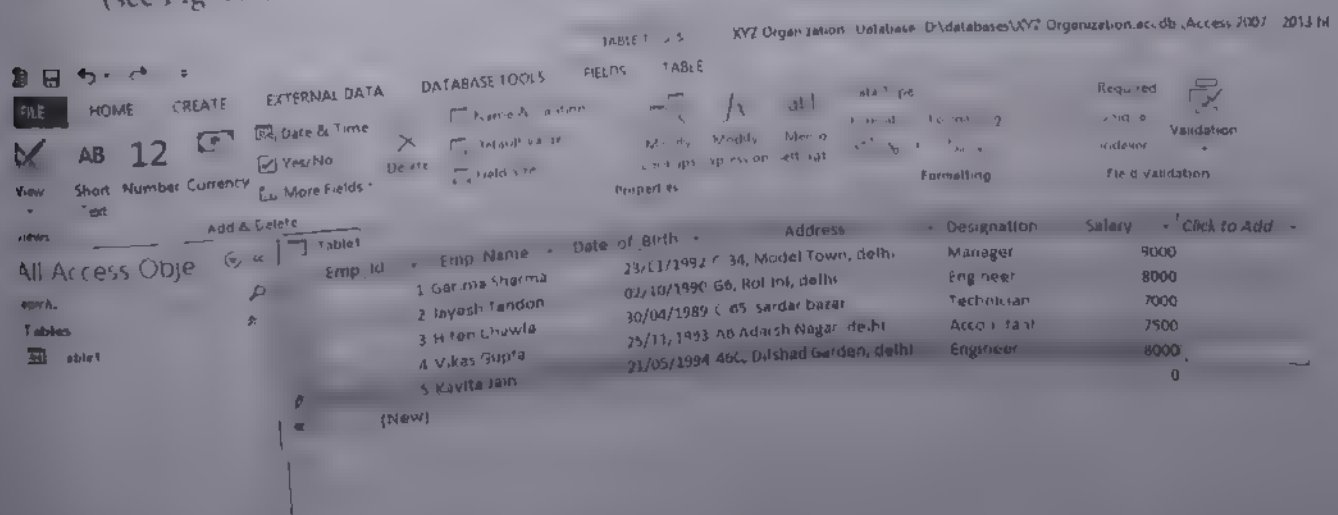


Figure 9.7 Entering data

Step 4: Click on the pencil icon to save the records. Close the table by clicking on the Close button.



- A field name can have 1 to 64 characters. It can have letters, numbers, and some special characters like underscore (_). A field name cannot have [], exclamation mark (!), or a period (.). A field name cannot begin with a blank space. It can be written in uppercase as well as lowercase.
- You can rename the field name by double clicking on it and then typing the new name. Or, you can even right click on a field's name (e.g., Field1) and from the menu choose Rename field. Note that we have renamed all the fields in our table.
- To rename the table, right click on Table1 in the Navigation pane and click on Rename. Enter the desired table name and click on OK.
- The new table name appears in the navigation pane as well. To change the data type of a particular field, click on the field header whose data type you want to change, a drop-down list appears from which you can select an appropriate type.

Adding or Deleting Fields

Step 1: Click the column to the right of which a new field is to be added.

Step 2: Click the data type required for the new field. You can also select the data type from the Add & Delete group of Field tab. A new field will be inserted (see Fig. 9.8).

Step 3: Similarly, to delete a field, click on the column header of the

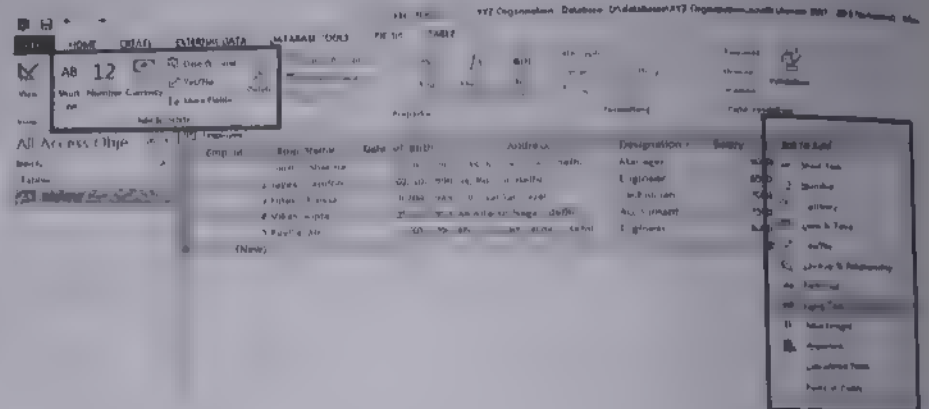


Figure 9.8 Adding a field

field to be deleted. From the menu, select Delete field. Another way to delete a field is to click on the field followed by clicking Delete button in the Add & Delete tab of Field tab, as shown in Fig. 9.9.

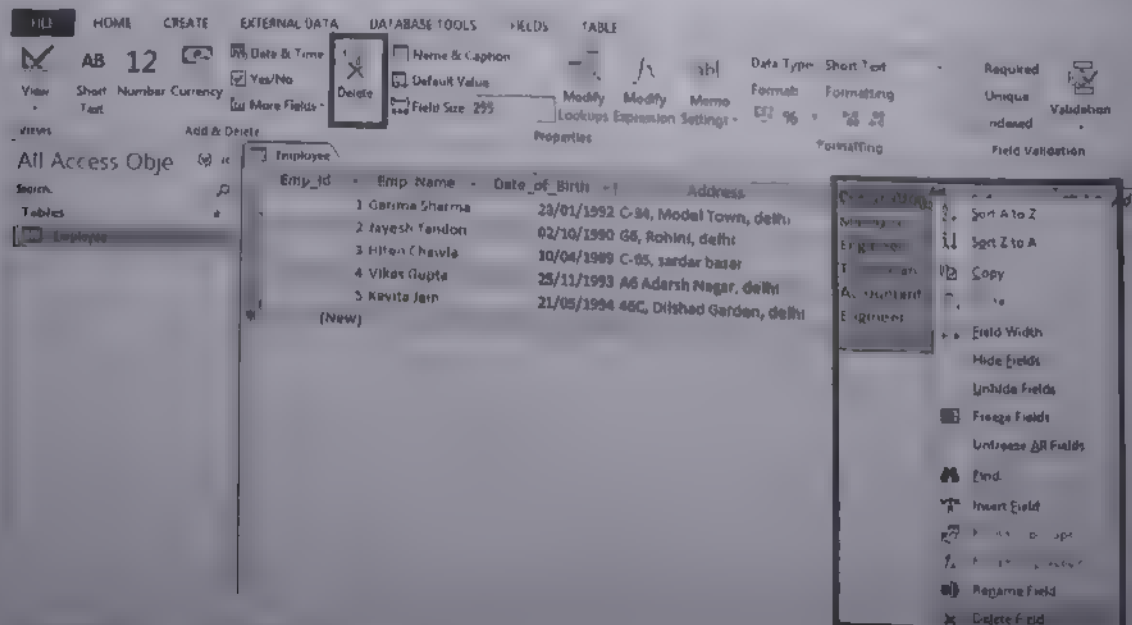


Figure 9.9 Deleting a field



- To move a field, simply double click on it and drag the field to left or right. Once the field is placed in the desired position, release the mouse.
- To insert a new field in between two fields, add it at the end and move it to the desired location.

9.4.2 Creating Tables in Design View

Open any previously created database. To add new tables in this database,

Step 1: Click on Create tab.

Step 2: In the Tables group, click on Table Design button to open the Design View (see Fig. 9.10).

Step 3: Set Dept No field as the primary key by right clicking on the field and selecting primary key from the menu. Alternatively, you can also select the field and then click on the primary key button in Tools group of Design tab (see Fig. 9.11). A key icon appears in the selector area to the left of the field's name.

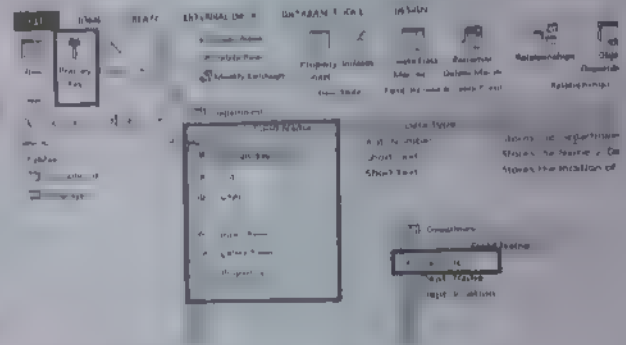
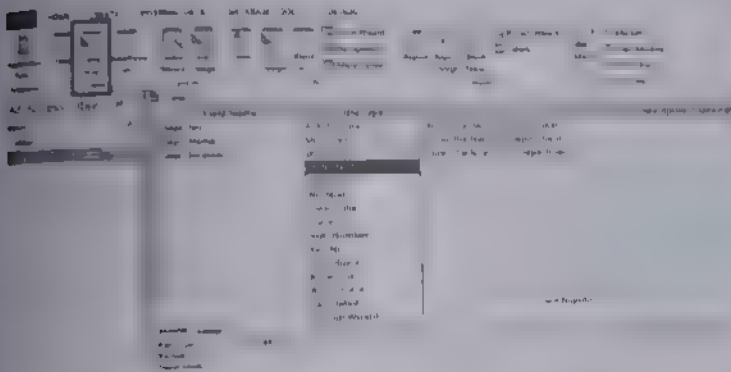


Figure 9.10 Creating a table in design view

Figure 9.11 Setting the primary key

Step 4: After adding the fields and setting their data types, save the newly created table and name it Department.



- If you try to save a table without specifying the primary key, then Access will automatically create a primary key field having AutoNumber data type
- If you define the primary key on a field after entering the data and that field has duplicate values, then an error message will be displayed. To remove the error message, you need to remove the duplicate values
- Primary key can be set as one particular field or as a combination of multiple fields. To set the primary key with multiple fields, hold down the Ctrl key and click on the Field Selector of each field
- To remove the primary key, click primary key in Tools group of the Design tab. The key icon disappears
- Once set, Access does not allow users to delete the Primary Key field. If you want to delete the field, first remove the primary key and then delete that field.

Example 9.2 Insert some records in the newly created Department table.

Solution: Switch from the Design View to Datasheet View from View button on the View group and insert the records as shown in Fig. 9.12.

9.4.3 Defining Relationship among Multiple Tables

MS Access uses table relationships to join tables when you need to use them in a database object.

When you design a database, you divide your information into tables, each of which has a primary key and then add foreign keys to related tables that reference those primary keys.

These foreign key–primary key pairings form the basis for table relationships and multi-table queries.

There are several reasons why you should create table relationships before you create other database objects, such as forms, queries, macros, and reports.

- To work with records from more than one table, you often must create a query that joins the tables. The query works by matching the values in the primary key field of the first table with a foreign key field in the second table.
- When you design a form or report, MS Access uses the information it gathers from the table relationships you have already defined to present you with informed choices and to prepopulate property settings with appropriate default values.

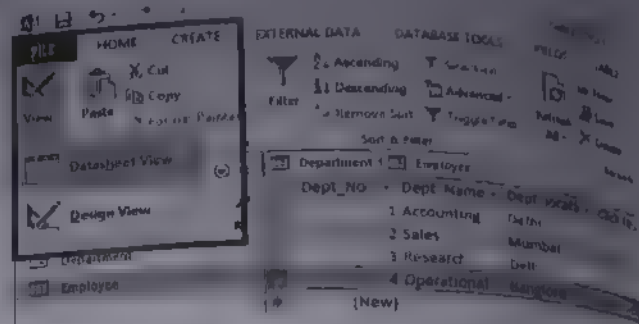


Figure 9.12 Inserting records

Example 9.3 Define a relationship between the Employee and Department tables by defining the primary key of the Department table as the foreign key in the Employee table.

Solution: Here, each department of the XYZ Organization can have more than one employee and each employee can work only in one department.

Step 1: Edit and add a new field in the Employee table in Design View as Dept_No (Foreign key) which refers to the Dept_No (primary key) of the Department table, as shown in Fig. 9.13.

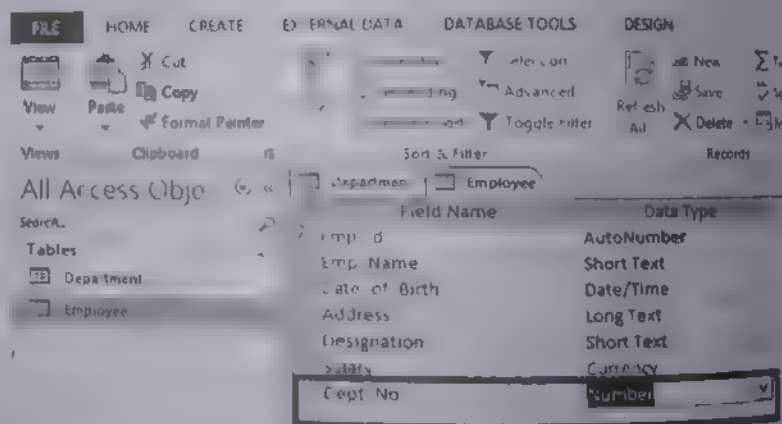


Figure 9.13 Step 1



Data type and field size of both the fields should match.

Step 2: Now go to Database Tools and select Relationships (Fig. 9.14).

Step 3: Click the Show table option and add Employee and Department tables. Close the Show table dialog box (see Fig. 9.15).

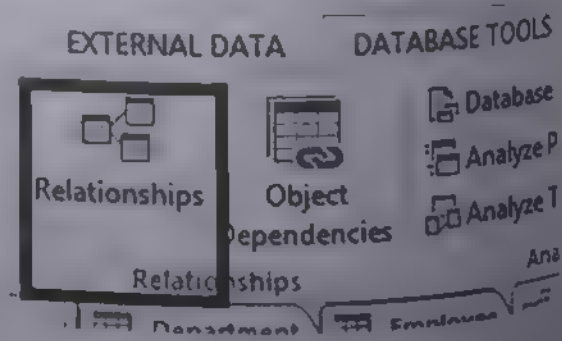


Figure 9.14 Step 2

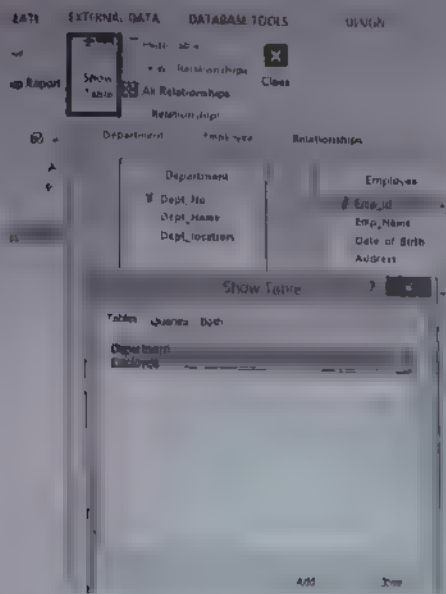


Figure 9.15 Step 3

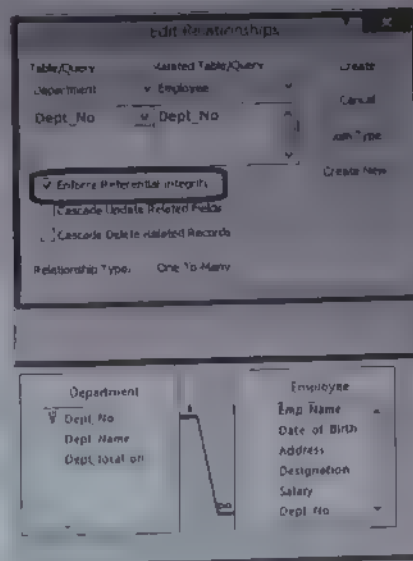


Figure 9.16 Step 4

Step 4: Click and hold Dept_No from Department and drag that all the way over to the Dept_No from Employee. Further, a relationships window pops up, as shown in Fig 9.16, when you release the mouse.



Do not forget to check the Enforce Referential Integrity option

Step 5: Click on the Create button.

Step 6: Insert the data for Dept_No field (foreign key) in the Employee table as shown in Fig 9.17

Emp_id	Emp_Name	Date_of_Birth	Address	Designation	Salary	Dept_No	Click to Add
1	Garima Sharma	23/01/1992	C-34, Model Town, delhi	Manager	9000	2	
2	Jayesh Tandon	02/10/1990	G6, Rohini, delhi	Engineer	8000	3	
3	Hiten Chawla	30/04/1989	C-65, sardar bazar	Technician	7000	4	
4	Vikas Gupta	25/11/1993	A6 Adarsh Nagar, delhi	Accountant	5500	1	
5	Kavita Jain	21/05/1994	46C, Dilshad Garden, delhi	Engineer	8000	5	

(a)

Emp_id	Emp_Name	Date_of_Birth	Address	Designation	Salary	Dept_No	Click to Add
1	Garima Sharma	23/01/1992	C-34, Model Town, delhi	Manager	9000	2	
2	Jayesh Tandon	02/10/1990	G6, Rohini, delhi	Engineer	8000	3	
3	Hiten Chawla	30/04/1989	C-65, sardar bazar	Technician	7000	4	
4	Vikas Gupta	25/11/1993	A6 Adarsh Nagar, delhi	Accountant	5500	1	
5	Kavita Jain	21/05/1994	46C, Dilshad Garden, delhi	Engineer	8000	5	



(b)

Figure 9.17 Step 6



- Referential integrity is a property of data stating that references within it are valid in the context of relational databases. It requires every value of one attribute (column) of a relation (table) to exist as a value of another attribute (column) in a different (or the same) relation (table).
- MS Access does not allow you to store any other value in Dept_No field of Employee table which is not there in Dept_No field of Department table, else it will show error.

9.5 QUERY OBJECT

An object provides a custom view of data from one or more tables. Queries are a way of searching for and compiling data from one or more tables.

Running a query is like asking a detailed question of your database. When you build a query in Access, you are defining specific search conditions to find exactly the data you want.

You can define queries to select, update, insert, or delete data. You can also define queries that create new tables from data in one or more existing tables.

9.5.1 Creating Queries using Query Wizard

Step 1: Click on Query Wizard in Queries group in the Create tab.

Step 2: A New Query dialog box opens up as shown in Fig. 9.18. Select Simple Query Wizard and click Ok.

Step 3: In the dialog box that is displayed, select the table that contains your data.

Step 4: The fields of the selected table will be displayed in the Available Fields Box. Add or remove the fields in the Selected Fields textbox using the appropriate buttons as shown in the figure.

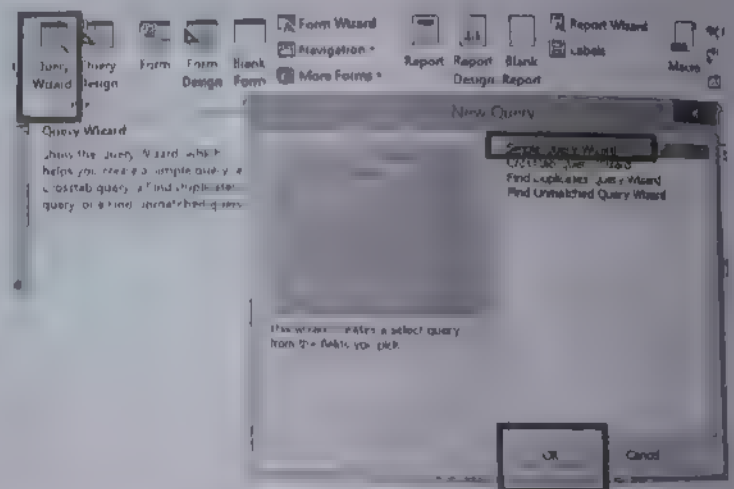


Figure 9.18 New query



- Query can be used to view, insert, delete records, or apply some calculations on the values given in one or more tables.
- Select the field and click on > and < button to add or remove it in/from the Selected Fields box respectively.
- To move all the fields from the Available list to Selected list, press the >> button. Use the << button for doing the reverse.

Example 9.4 Create a query to select Emp Name, Designation, and Salary from Employee table
Solution: This example follows the aforementioned steps 1–4. Add Emp_Name, Designation, and Salary fields from Available list to Selected list.

Step 5: Click on Next.

Step 6: From the dialog box, select either Detail or Summary as required and click on Next, as shown in Fig. 9.19

Step 7: In the dialog box, type an appropriate name for the query. Select open the query to view information and finally click on Finish (see Fig. 9.20). The query result will be shown in Datasheet View.

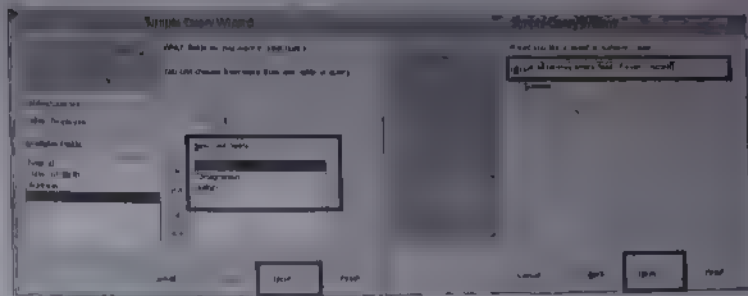


Figure 9.19 Example 9.4

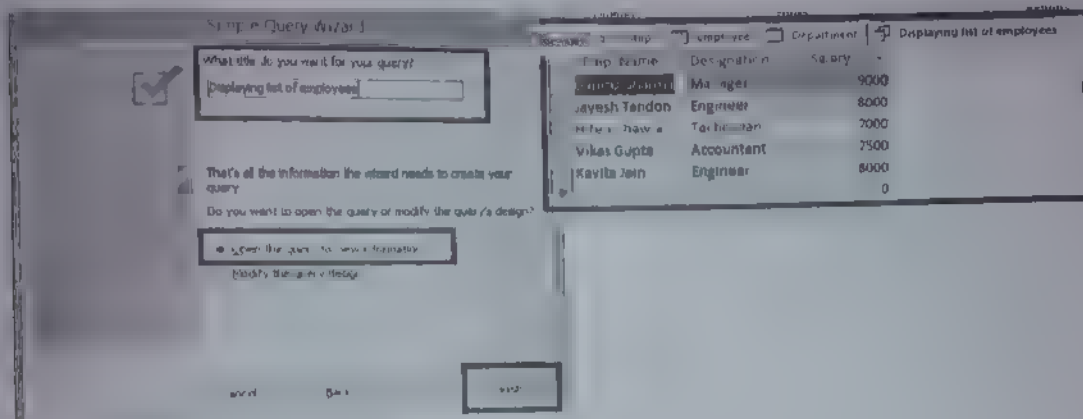


Figure 9.20 Opening the query

9.5.2 Creating Queries using Query Design

Step 1: Click on Query Design in Queries group of the Create tab

Step 2: A Show Table dialog box will appear on the screen as shown in Fig. 9.21. Select the Table and click on Add.

Step 3: Close the Show Table Dialog box and select the fields to be included in the query by double clicking on the field's name



Double click on * to select all the fields of the table in the Grid and Query as well.

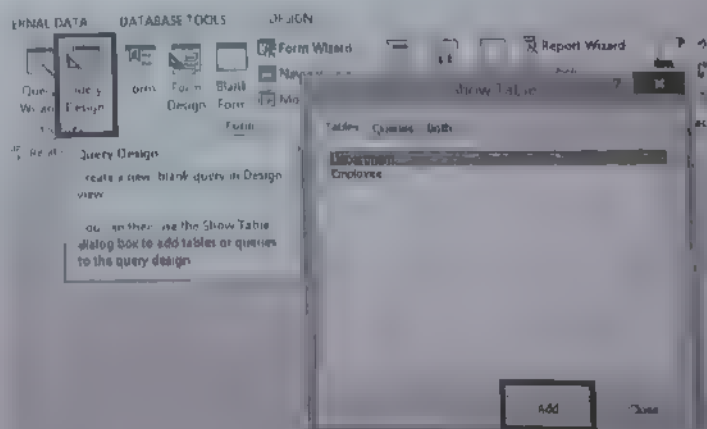


Figure 9.21 Show table dialog box

Example 9.5 Create a select query to display a list of Dept_Name and Dept_location from Department table.

Solution: Select Department table from the Show Table dialog box and select Dept Name and Dept_location fields as shown in Fig. 9.22.

Save the query and click on Run button in the Results group of Design tab.

Example 9.6 Create a conditional query that displays Employees details of those employees whose salary is greater than 7500.

Solution: Step 1: Click on Query Design in Queries group of the Create tab.

Step 2: Add Employee table from Show Table dialog box and select Emp Name, Dept No, Designation, and Salary fields.

Step 3: Go to the criteria row of salary and type >7500.

Step 4: Save and run the query (see Fig. 9.23).

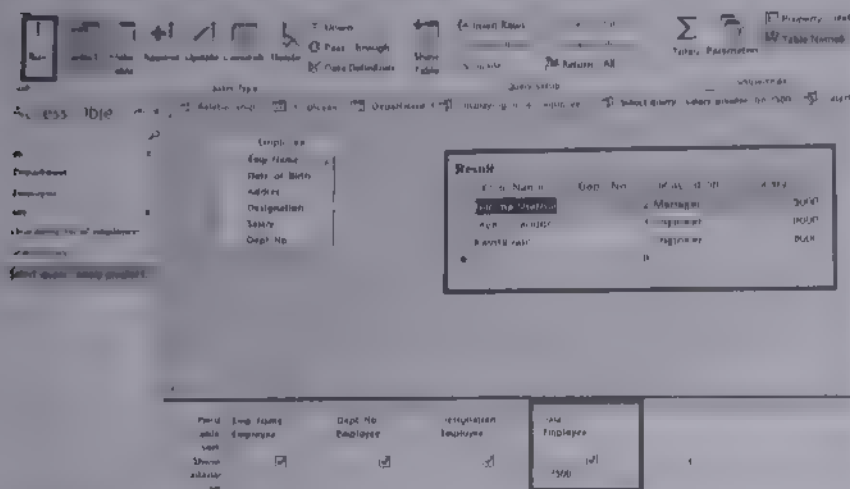


Figure 9.23 Example 9.6

Step 4: Save and run the query (see Fig. 9.24).

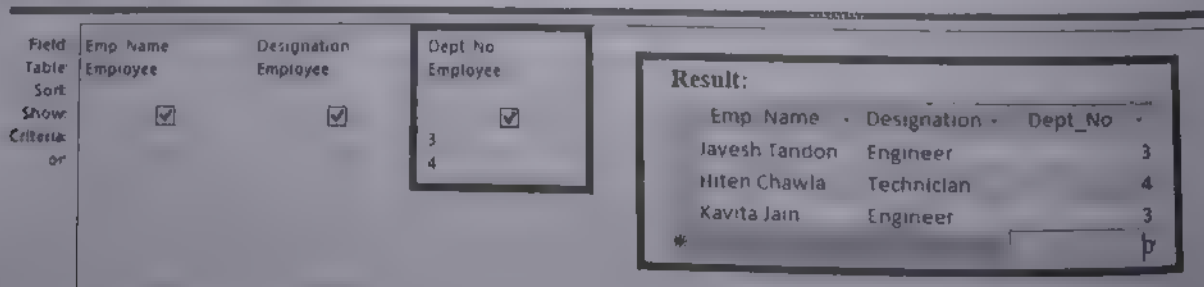


Figure 9.24 Example 9.7

Example 9.8 Create a conditional query that displays Employees details of those employees whose Designation is 'Manager' and Salary is greater than 7500.

Solution: Step 1: Click on Query Design in Queries group of the Create tab.

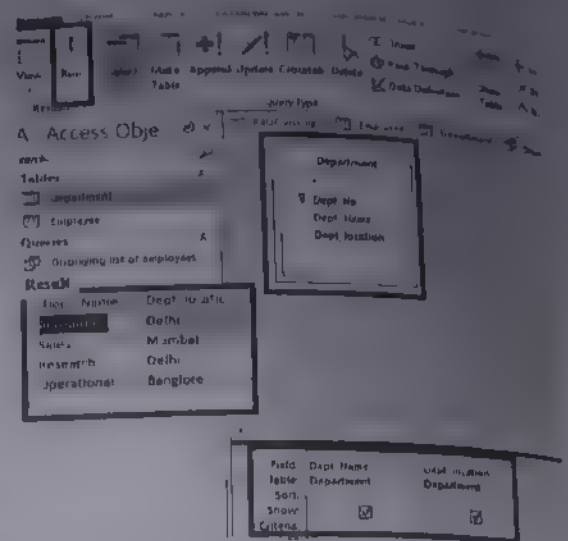


Figure 9.22 Selection of fields

Example 9.7 Create a conditional query that displays Employees details of those employees who work in either department number 3 or 4.

Solution: Step 1: Click on Query Design in Queries group of the Create tab.

Step 2: Add Employee table from Show Table dialog box and select Emp Name, Dept No, and Designation fields.

Step 3: Go to the criteria row of Dept No, type 3 and type 4 in Or row to select either of the Dept No.

Step 2: Add Employee table from Show Table dialog box and select Emp_Name, Dept_No, Designation, and Salary fields.

Step 3: Go to the criteria row of Designation, type Manager and in criteria row of salary type >7500.

Step 4: Save and run the query (see Fig 9.25).

Field:	Emp_Name	Designation	Dept_No	Salary
Table:	Employee	Employee	Employee	Employee
Sort:				
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Criteria:		Manager		>7500

Emp_Name	Designation	Dept_No	Salary
Garima Sharma	Manager	2	9000

Figure 9.25 Example 9.8

Example 9.9 Create an update query to update salary of employee whose Emp_id is 3.

Solution: Step 1: Click on Query Design in Queries group of the Create tab.

Step 2: Add Employee table from Show Table dialog box.

Step 3: On the Design tab, in the Query Type group, click Update and double click on the field in which you want to update the value.

Step 4: In the Update row of the Design grid, enter the updated value and in Criteria row add [Employee] [Emp_id] = 3 and run the query. This will display the confirmation message as shown in Fig 9.26.

Step 5: Click Yes and go to Datasheet View and you will see that the record is now updated.

You are about to update 1 row(s).
 Are you sure you want to update these records?
 Yes No

Field:	Salary
Table:	Employee
Update to:	7500
Criteria:	[Employee] [Emp_id]=3

Emp_id	Emp_Name	Date of Birth	Address	Designation	Salary	Dept_No
1	Garima Sharma	23/01/1992	C-34 Model Town, delhi	Manager	9000	2
2	Jayesh Tandon	02/10/1990	G-6, Rohini, delhi	Engineer	8000	3
3	Hiten Chawla	30/04/1989	C-65, Sardar Bazar	Technician	7500	4
4	Vikas Gupta	25/11/1993	A-6 Adarsh Nagar, delhi	Accountant	7500	1
5	Kavita Jain	21/05/1994	40C, Dilshad Garden, delhi	Engineer	8000	3
					0	0

Figure 9.26 Example 9.9

Example 9.10 Create a delete query to delete an employee whose Emp_id is 2.

Solution: Step 1: Click on Query Design in Queries group of the Create tab.

Step 2: Add Employee table from Show Table dialog box.

Step 3: On the Design tab, in the Query Type group, click Delete and double-click on the Emp_id.

Step 4: In the Criteria row of the Design grid, type 2 and run the query. This will display the confirmation message shown in Fig. 9.27.

Step 5: Click Yes and go to Datasheet View and you will see that record is now deleted.

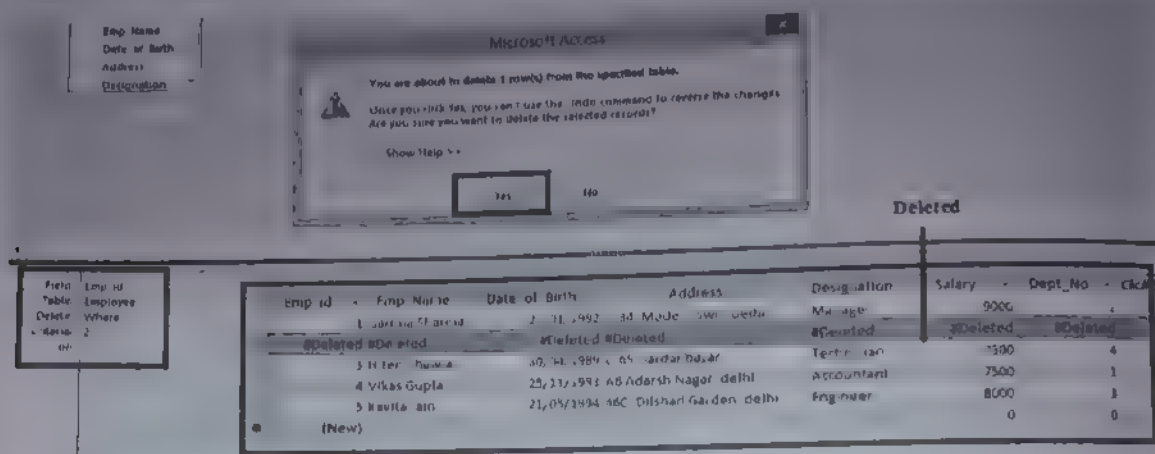


Figure 9.27 Example 9.10

Example 9.11 Create an append query to append all the records of Employee table to Temp_Emp table.

Solution: Step 1: Create Temp_Emp in Design View with all fields as that in Employee table (see Fig. 9.28).

Step 2: Click on Query Design in Queries group of the Create tab.

Field Name	Data Type
Emp_id	AutoNumber
Emp_Name	Short Text
Date of Birth	Date/Time
Address	Short Text
Designation	Short Text
Salary	Number
Dept_No	Number

Figure 9.28 Example 9.11

Step 3: Add Employee table from Show Table dialog box and Click on * to select all fields of Employee table (see Fig. 9.29).

Step 4: On the Design tab, in the Query Type group, click Append and select Temp_Emp table in Append dialog box.

Step 5: Run the query. This will display the confirmation message.

Step 6: Click Yes and go to Datasheet View and you will see that the records are now appended (Fig. 9.30).

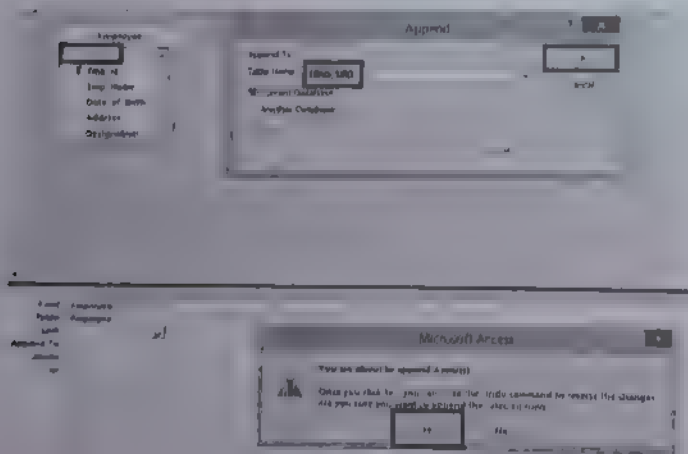


Figure 9.29 Adding employee table

Emp_id	Emp_Name	Date_of_Bir	Address	Designation	salary	Dept_No	Click
1	Garena Shorm	23/01/1992	C-34, Model T	Manager	9000	2	
3	Hiten Chawla	30/04/1989	C-65, Sardar b	Technician	7300	4	
4	Vikas Gupta	25/11/1993	A-6 Adarsh Nag	Accountant	7500	1	
5	Kavita Jain	21/05/1994	A-6C, Dilshad G	Engineer	8000	3	
(New)					0	0	

Figure 9.30 Viewing the appended records

9.6 FORM OBJECT

Form is an object in a desktop database designed primarily for data input or display or for control of application execution. You use forms to customize the presentation of data that your application extracts from queries or tables.

Forms are used for entering, modifying, and viewing records. The reason forms are used so often is that they are an easy way to guide people toward entering data correctly.

When you enter information into a form in Access, the data goes exactly where the database designer wants it to go in one or more related tables.

9.6.1 Creating Forms using Form Command

Step 1: Select a table by double clicking on it in the Navigation Pane.

Step 2: Click Form in the Forms group of Create tab. The following screen shown in Fig. 9.31 appears.

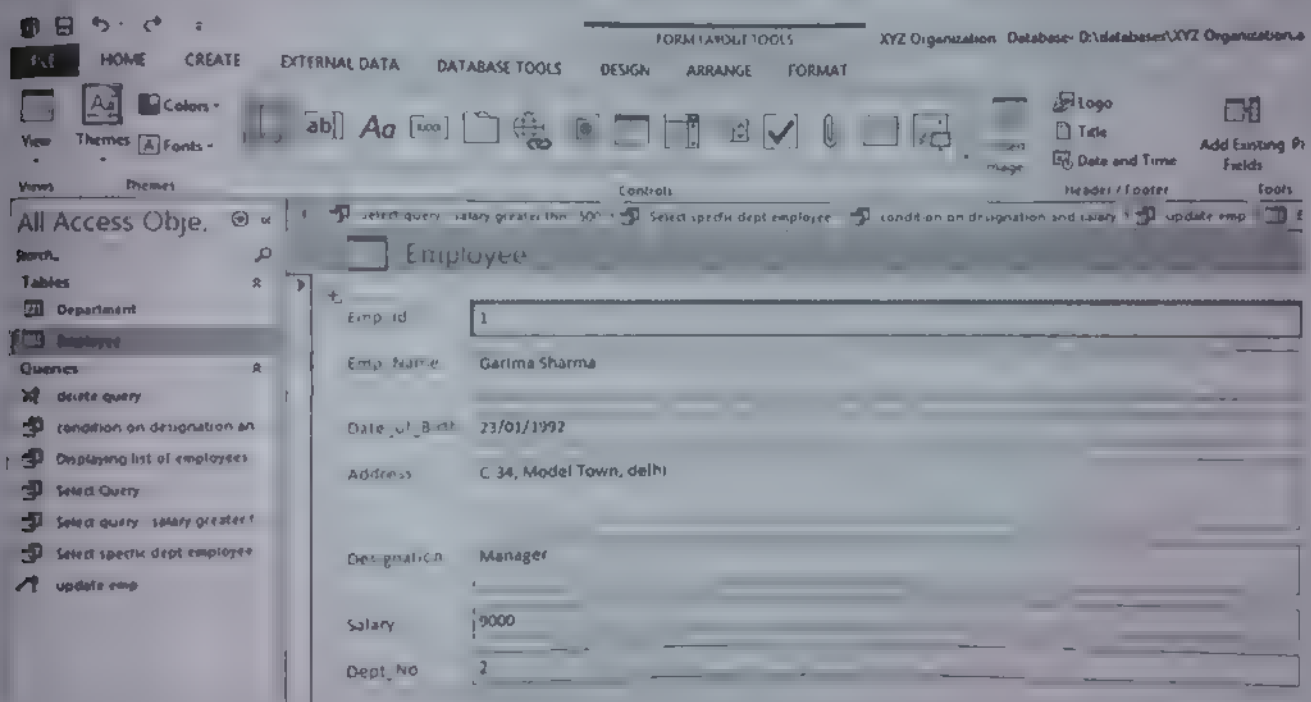


Figure 9.31 Creating a form

It has all the fields used in the table. Notice that the Design, Arrange, and Format tools appear under Form Layout Tools tab.

Step 3: Click on Logo in Header/Footer group. Select any picture from the Insert Picture dialog box that appears and click on OK.

Step 4: Click on Title in Header/Footer group and type an appropriate title for the form.

Step 5: Click on Date and Time in Header/Footer group. Select any format and click on OK.

Step 6: To change the theme, click on Themes in Themes group and select one of your choice.

To add new employee information, go to the end of these records and then after 4 records you will see a blank form where you can begin entering the new employee's information (Fig. 9.32).

Step 7: Click on the Format tab under Form Layout Tools. Now you can add colours to the labels, change the font colour and size, alignment and background colour etc., by clicking on the desired option.

Step 8: To save the form, click on File tab and select Save option.

Step 9: Enter a suitable form name and click on Ok.

After saving the form, you can see that the tables, queries, and forms in your database can be seen in the Navigation Pane. To open them you can simply double click on them.

Figure 9.32 Adding new employee information

9.6.2 Creating Forms using Split Form

Step 1: Open any previously created table by double clicking on its name in the Navigation Pane.

Step 2: Click on More Forms in the Forms group of Create tab and Select Split Forms. The following screen, shown in Fig. 9.33 appears.

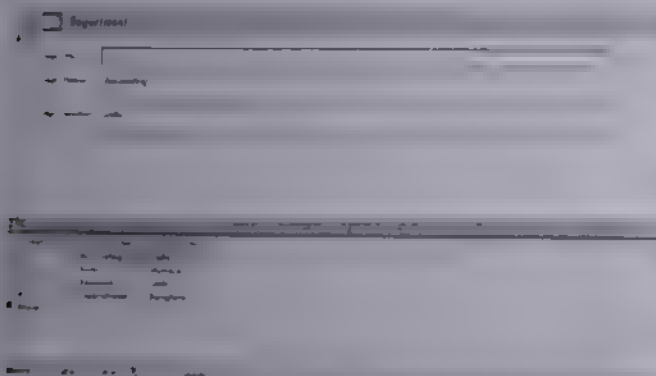


Figure 9.33 Creating a form using split form

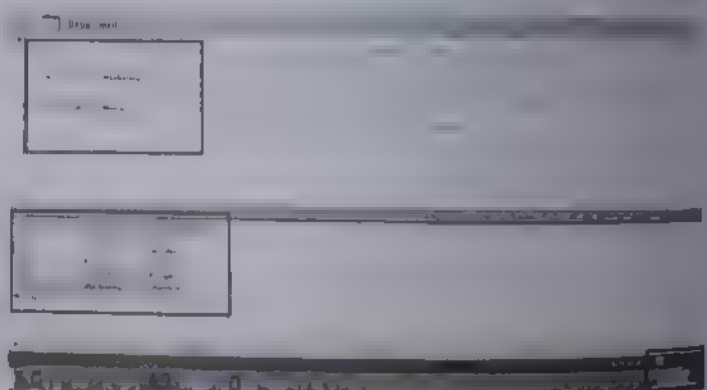


Figure 9.34 Form view

Step 3: Click the Form View button at the bottom right end of status bar to enter or change data, as shown in Fig. 9.34.

9.6.3 Creating Forms using Form Wizard

Step 1: Open any previously created table by double clicking on its name in the Navigation Pane.

Step 2: Click on Form Wizard in the Forms group of Create tab.

Step 3: From the Dialog Box, select the required table from the available Tables or Queries in the database.

Step 4: Select the fields required in the Form and click on Next, as shown in Fig. 9.35

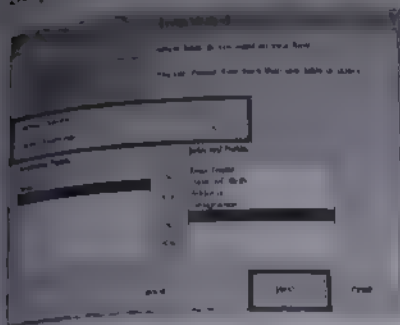


Figure 9.35 Form wizard

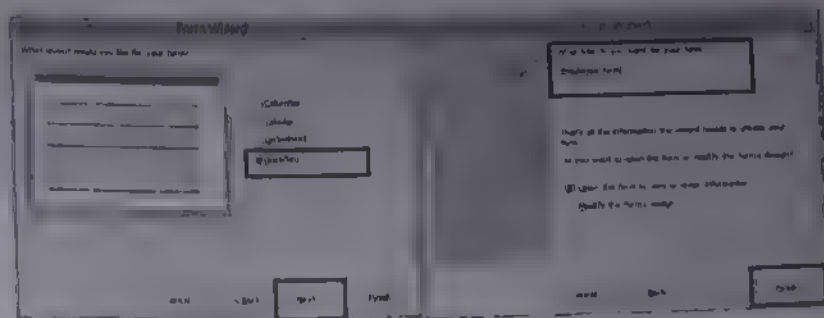


Figure 9.36 Next steps in form wizard

Step 5: Select the form layout and click on Next.

Step 6: Give a suitable name to the form and click on Open the form to view or enter information, as shown in Fig. 9.36.

Step 7: Click Finish.

9.7 REPORT OBJECT

Report is an object in desktop databases designed for formatting, calculating, printing, and summarizing the selected data.

You can view a report on your screen before you print it. Anything you plan to print deserves a report, whether it is a list of names and addresses, a financial summary for a period, or a set of mailing labels.

Reports are useful because they allow you to present components of your database in an easy-to-read format.

Access offers you the ability to create a report from any table or query.

9.7.1 Creating Reports using Report Command

Step 1: Double click on any table to open it.

Step 2: Click on Reports in the Reports group of Create tab.

Step 3: Select the layout view to format the report.

Step 4: Click on the Print Preview button at the bottom right corner of status bar to see how the report looks. See Fig. 9.37.

Emp id	Emp name	Date of Birth	Address	Designation	Salary	Dept No
1	Garna Sharma	23/0 /1992	C-34, Model Town, delhi	Manager	9000	2
2	Hiten Chowd	30/04/1989	C-65, sardar bazar	Technician	7300	4
3	Vikas Gupta	25/ /1993	A6 Adarsh Nagar delhi	Accountant	7500	1
4	Kavita Jain	21/05/1994	46C, Dishaad Garden, delh	Engineer	8000	3
					31800	

Figure 9.37 Report created using report command

9.7.2 Creating Reports using Report Wizard

Step 1: Double click on any table to open it and click on Report Wizard in the Reports group of Create tab

Step 2: Select the table in Table/Queries and the fields required in the report and click on Next.

Step 3: Select the grouping level in report and click on Next (see Fig. 9.38).

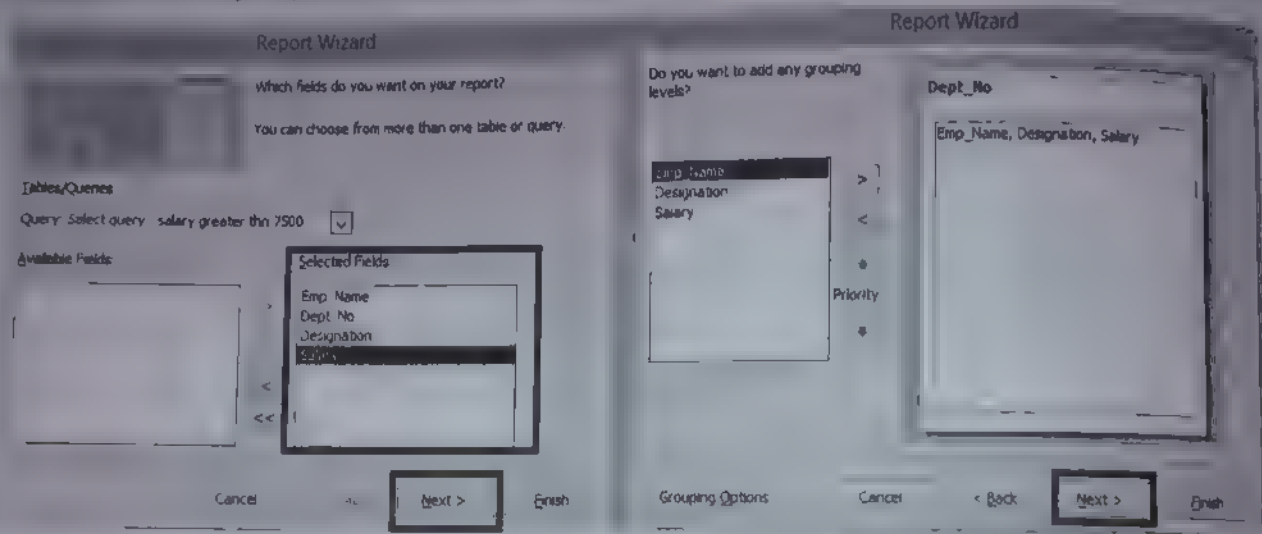


Figure 9.38 Report wizard

Step 4: From the dialog box, you can select the sorting order and summary options.

Step 5: Click on Next. From the dialog box, select the layout of the report. Also select either portrait or landscape in orientation section and click next (see Fig. 9.39).

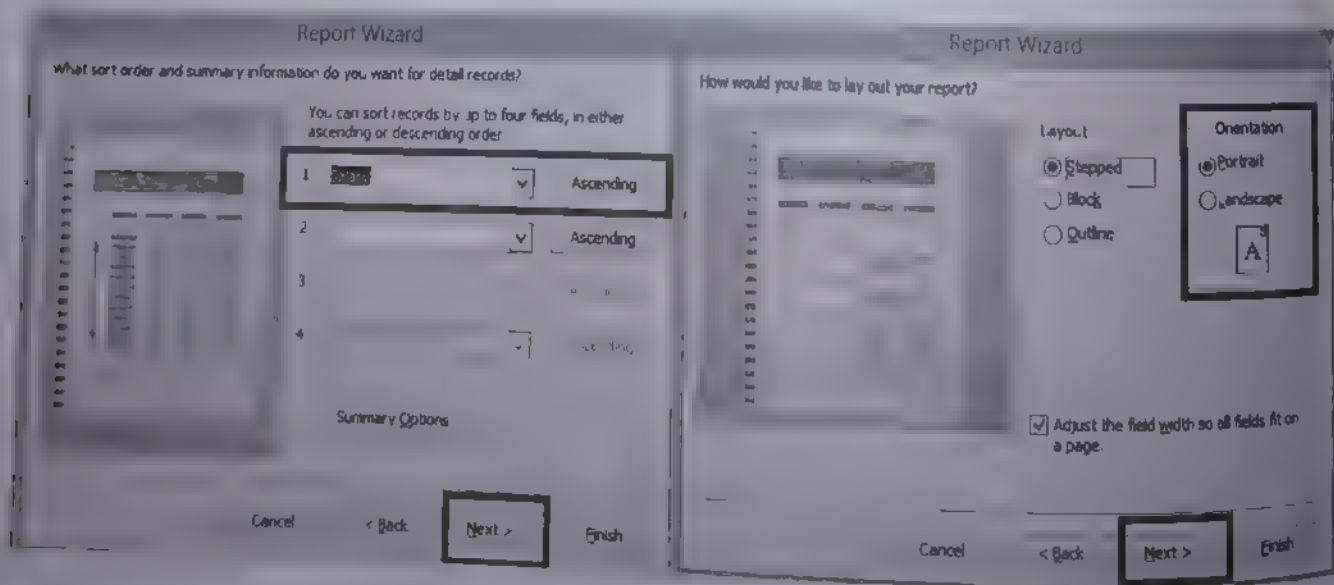


Figure 9.39 Next steps in report wizard

Step 6: Enter a title for the report. Select Preview the report and click on Finish as shown in Fig. 9.40.

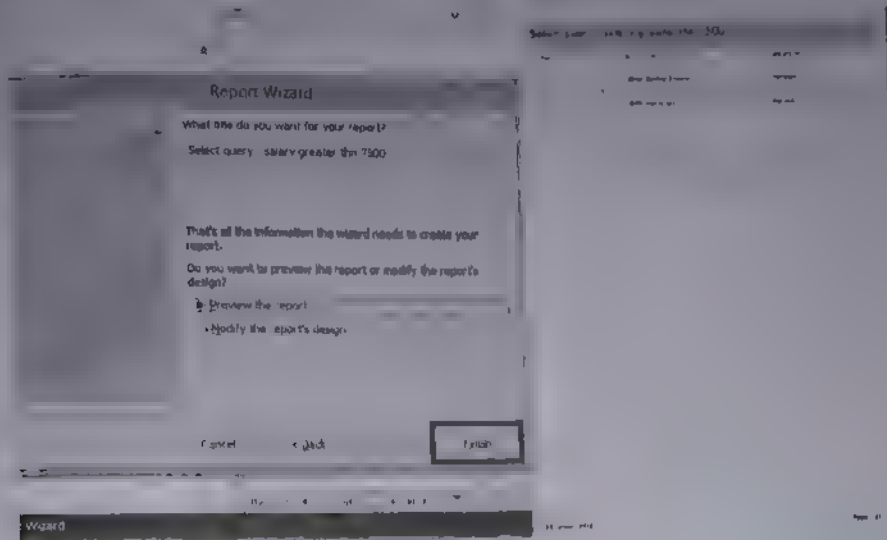


Figure 9.40 Steps to preview the report

The report will be displayed on the screen. You can print this report using the Print option in the File menu.

9.8 INTRODUCTION TO SQL THROUGH BASIC COMMANDS

Structured query language (SQL) is a computer language for storing, manipulating, and retrieving the data stored in a relational database. SQL is the standard language for a relational database system. All the RDMS such as MySQL, MS Access, Oracle, Sybase, and SQL Server, use SQL as their standard database language.

Whenever you create a query in Query Design, Access automatically creates the SQL query for you. This actually retrieves data from the tables. We can see how a query is created in SQL when we create it in Query Design.

Example 9.11 Show the steps to see a query which is created in SQL for a query (to select Emp Name, Designation, and Salary from Employee table) which is already designed

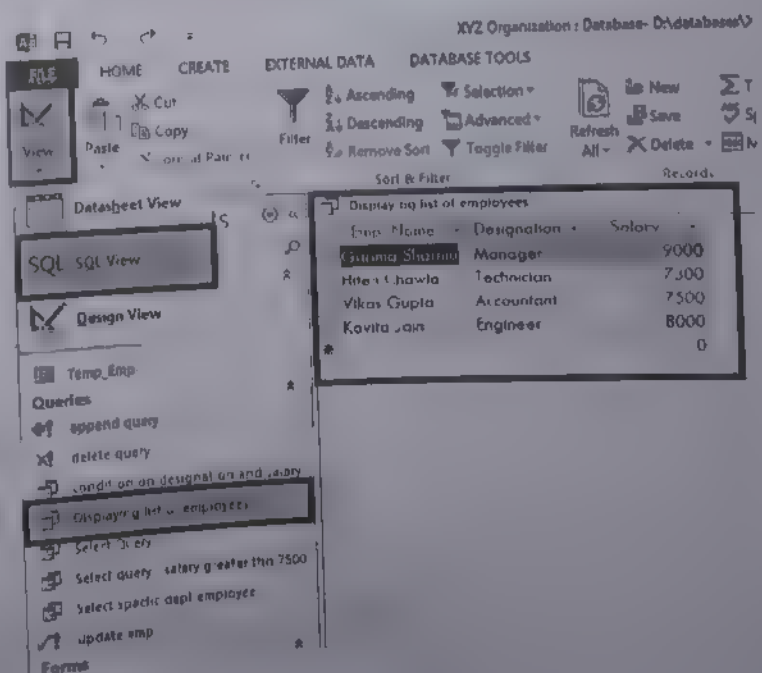


Figure 9.41 Opening a query

Solution: Step 1: Open a query from the Navigation Pane as shown in Fig. 9.41.

Step 2: Now select the SQL View from the View Menu in Home tab. You will see the SQL of your query (Fig. 9.42).

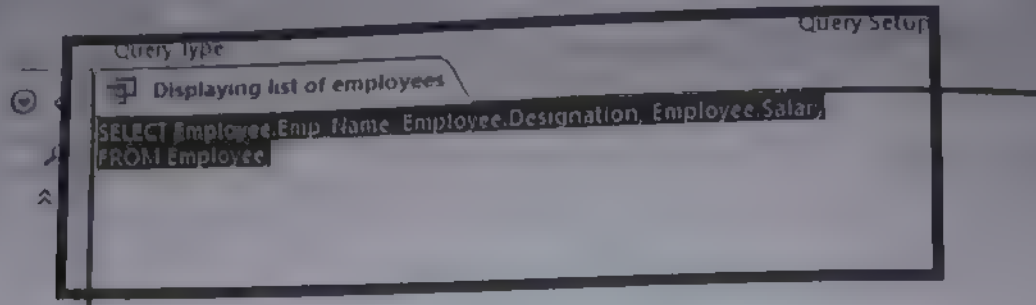
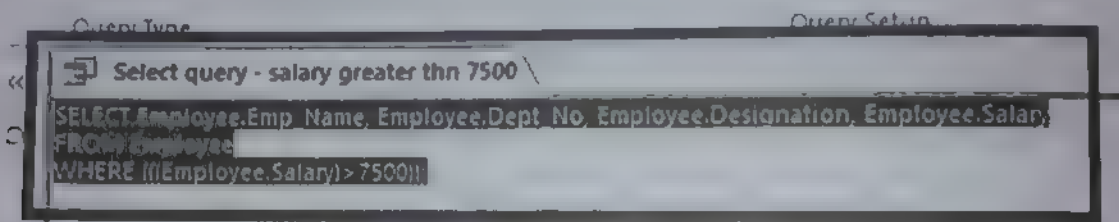


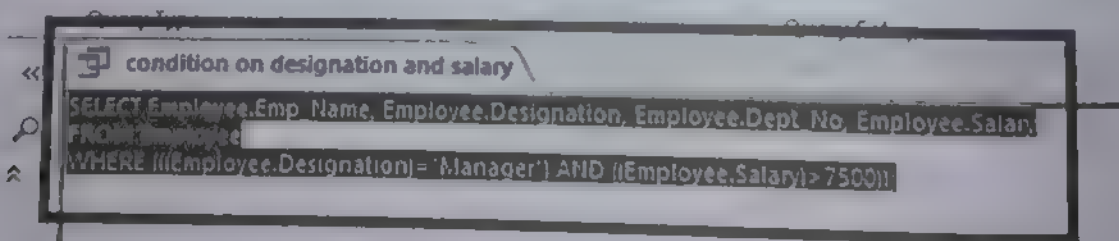
Figure 9.42 SQL of a query

Figures 9.43(a) (c) show the SQL queries generated for various queries. Fig. 9.43(a) shows the SQL query of the conditional query that displays the details of those employees whose salary is greater than 7500.



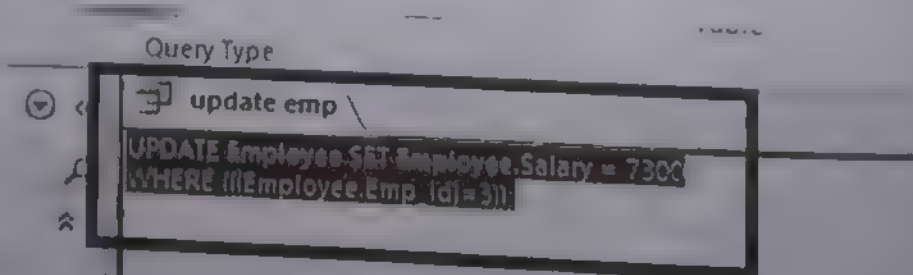
(a)

Figure 9.43(b) shows the SQL query generated for a conditional query that displays the details of those employees whose Designation is 'Manager' and Salary is greater than 7500.



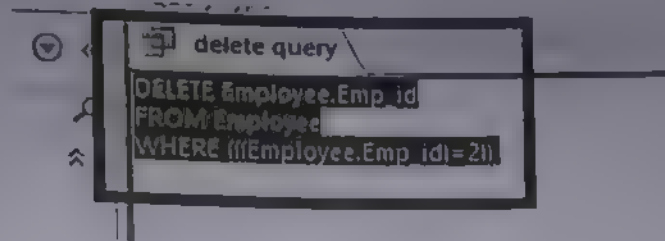
(b)

Figure 9.43(c) shows the SQL query generated for an update query to update salary of employee whose Emp_id is 3.



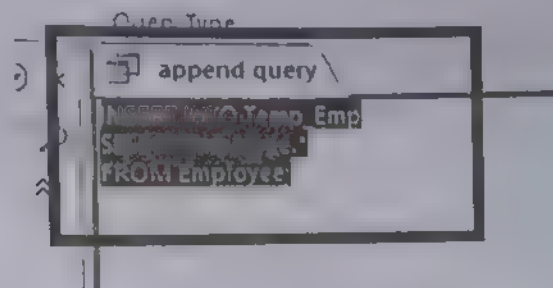
(c)

Figure 9.43(d) shows the SQL query generated for a delete query to delete an employee whose Emp_id is 2



(d)

Figure 9.43(e) shows the SQL query generated for an append query to append all the records of Employee table to Temp Emp table



(e)

Fig. 9.43 SQL queries (a) Salary > 7500 (b) Designation = Manager and Salary > 7500 (c) Updating details of Emp_id = 3 (d) Deleting Emp_id = 2 (e) Appending records

9.9 APPLICATION OF DBMS

Database management systems find application in a large number of fields today. Some of them are discussed in this section.

Accounting It is a systematic process of identifying, recording, measuring, classifying, verifying, summarizing, interpreting, and communicating financial information. In Access, Accounting can be achieved through the report object which is used for formatting, calculating, printing, and summarizing selected data.

Human resource management (HRM) It is the practice of recruiting, hiring, deploying, and managing an organization's employees.

In DBMS, HRM is done using the Employee and Department tables used in this chapter. All the employees of the organization, their salaries, and the department they work in are managed using these two tables.

Inventory This refers to the raw materials, work-in process products, and finished goods that are considered to be the portion of a business's assets that are ready or will be ready for sale.

Example 9.12 Create an RDBMS to help an organization manage its inventory.

Solution: Create two tables Inventory and Order in Design View with fields as shown in Fig 9.44(a) and (b)

Inventory		Order	
Field Name	Data Type	Field Name	Data Type
Item_ID	AutoNumber	Order_ID	AutoNumber
Item_Description	Short Text	Sell_Price	Number
Item_Price	Number	Date_Ordered	Date/Time
Item_Stock	Number	Qty_Ordered	Number
		Item_ID	Number

(a) (b)

Figure 9.44 Field details of tables to be created (a) Inventory (b) Order

Create a relationship between the two tables by defining the primary key of Inventory table as the foreign key in Order table, as shown in Fig 9.44. Do not forget to check 'Enforce Referential Integrity'

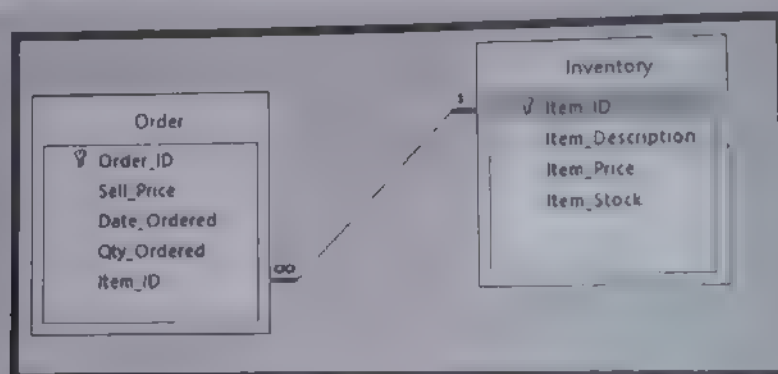


Figure 9.45 Creating a relationship between tables

With the help of these tables, the organization can keep track of goods or assets and even manage the order details for these products.

Summary

- Microsoft Access allows you to link related information easily. It can also import or link directly to data stored in other applications and databases.
- Databases in Access are composed of tables, queries, forms, and report objects.
- In Access, a field can store different types of data.
- A field name can have 1 to 64 characters.
- An auto-number field stores an integer whose value increases or decreases automatically as new records are added or deleted.
- Access uses table relationships to join tables when you need to use them in a database object.
- You can create a query using Query Wizard and Query Design.
- Forms are used often as they provide an easy way to guide people in entering data correctly.
- Access offers you the ability to create a report from any table or query.

Glossary

Microsoft Access A DBMS from Microsoft that combines the relational Microsoft Jet Database Engine with a graphical user interface and software development tools

Database An organized collection of data

Table An object that is used to define and store data

Primary key Used to uniquely identify each record in a table

Foreign key A field (or collection of fields) in one table that uniquely identifies a row of another table or the same table

Query An object that provides a custom view of data from one or more tables

Form An object in a desktop database designed primarily for data input or display or for control of application execution

Report An object in desktop databases designed for formatting, calculating, printing, and summarizing selected data

Multiple-choice Questions

1. Queries in Access can be used _____.
 - (a) to view, change, and analyse data in different ways
 - (b) as a source of records for forms and reports
 - (c) both (a) and (b)
 - (d) none of these
2. To create queries in Access, you can _____.
 - (a) drag and drop fields on query builders
 - (b) type the SQL command in SQL View
 - (c) use Query Wizard or Design View
 - (d) all of these
3. Query Design window has two parts. The upper part shows _____.
 - (a) name of fields, field type, and size
 - (b) tables with fields and relationships between tables
 - (c) criteria
 - (d) sorting check boxes
4. What does the show check box in Query Design window indicate?
 - (a) It indicates whether the field is to be used or not.
 - (b) It indicates whether the field is to be displayed in query result or not.
 - (c) It indicates whether the field names are to be displayed in query result or not.
 - (d) None of these
5. What is a form in MS Access?
 - (a) It is a printed page where users will write their data to fill it up.
 - (b) It is an input screen designed to make the viewing and entering of data easier.
 - (c) This is an important part of database used by analysts to draw conclusions.
 - (d) All of these
6. Which of the following statements is true?
 - (a) Reports can be used to retrieve data from tables and calculate.
 - (b) Queries can be printed in a well formatted manner and presented as the information.
 - (c) Queries can include calculated fields that do not exist in the table.
 - (d) Reports and forms are similar, but forms are used to print and reports are used to display on screen only.
7. Two tables can be linked with relationship so that the data integrity can be enforced. Where can you find the relationship command?
 - (a) File menu

- (b) Home menu
 - (c) Database
 - (d) Fields menu
8. To create a relationship between two tables _____.
- (a) drag the primary key of a table into foreign key of another table
 - (b) drag the foreign key of a table into the primary key of another table
 - (c) drag any field from parent table and drop on child table
 - (d) any of these can be done to create a relationship
9. What happens when you release the mouse pointer after you drop the primary key of a table into the foreign key of another table?
- (a) A relationship is created
 - (b) Edit relationship dialog box appears
 - (c) Error occurs
 - (d) Nothing happens
10. What do you mean by one to many relationship between Student and Class table?
- (a) One student can have many classes
 - (b) One class may have many students
 - (c) Many classes may have many students
 - (d) Many students may have many classes
11. Referential integrity means:
- (a) do not enter a value in the foreign key field of a child table if that value does not exist in the primary key of the parent table
 - (b) do not enter a value in the primary key field of child table if that value does not exist in the primary key of the parent table
 - (c) do not enter a value in the foreign key field of a parent table if that value does not exist in the primary key of the child table
 - (d) all of these
12. We can remove a relationship defined between two tables _____.
- (a) from Edit menu, choosing Delete Relationship
 - (b) by selecting the relationship line and pressing Delete
 - (c) by choosing Delete option from Relationship menu
 - (d) all of these
13. If you need to edit a relationship, you should _____.
- (a) right click the relationship line and choose Edit Relationship
 - (b) double click the relationship line
 - (c) both (a) and (b)
 - (d) none of these
14. If you write criteria values vertically (one in a row), it will mean _____.
- (a) OR condition
 - (b) AND condition
 - (c) NOT condition
 - (d) none of these
15. To achieve AND effect when you are entering criteria in Query Design window, _____.
- (a) write the criteria values vertically one in a row
 - (b) write the criteria values horizontally
 - (c) write the criteria values in same field separated with AND
 - (d) write the criteria values in same field separated with &
16. When a picture or other graphic image is placed in the report header section, it will appear ____.
- (a) once in the beginning of the report
 - (b) at the top of every page
 - (c) after every record break
 - (d) on the first and last pages of the report
17. _____ data type allows alphanumeric characters and special symbols.
- (a) Text
 - (b) Memo
 - (c) Auto number
 - (d) None of these
18. In a database table, the category of information is called ____.
- (a) tuple
 - (b) field
 - (c) record

- (d) all of these
19. To create a new table, in which method you do not need to specify the field type and size?
- (a) Create table in Design View
 - (b) Create table using wizard
 - (c) Create table by entering data
 - (d) All of these
20. Which of the following is not a database object?
- (a) Tables
 - (b) Queries
 - (c) Relationships
 - (d) Reports
21. The third stage in designing a database is when we analyse our tables more closely and create a _____ between tables
- (a) relationship
 - (b) join
 - (c) query
 - (d) none of these
22. The size of the Yes/No field is always _____
- (a) 1 bit
 - (b) 1 byte
 - (c) 1 character
 - (d) 1 GB
23. _____ is the stage in database design where one gathers and lists all the necessary fields for the database project
- (a) Data definition
 - (b) Data refinement
 - (c) Establishing relationship
 - (d) None of these
24. The size of a field with number data type cannot be _____
- (a) 2
 - (b) 4
 - (c) 8
 - (d) 16
25. _____ key uniquely identifies each record
- (a) Primary key
 - (b) Key record
 - (c) Unique key
 - (d) Field name
26. A database language concerned with the definition of the whole database structure and schema is _____.
- (a) DCL
 - (b) DML
 - (c) DDL
 - (d) all of these
27. When creating a new table, which method can be used to choose fields from standard databases and tables?
- (a) Create table in Design View
 - (b) Create table using wizard
 - (c) Create table by entering data
 - (d) None of these
28. Which field type will you select when creating a new table if you require to enter long text in that field?
- (a) Text
 - (b) Memo
 - (c) Currency
 - (d) Hyperlink
29. In table design view, what is the first column of buttons used for?
- (a) To indicate primary key
 - (b) To indicate current row
 - (c) Both of these
 - (d) None of these
30. Which option allows you to build a new table by entering data directly into the datasheet?
- (a) Datasheet View
 - (b) Design View
 - (c) Link table
 - (d) None of these
31. Which is not a view to display a table in Access?
- (a) Datasheet View
 - (b) Design View
 - (c) Pivot table and pivot chart view
 - (d) None of these
32. Which of the following database objects holds data?
- (a) Forms
 - (b) Reports

- (c) Queries
(d) Tables
33. A _____ enables you to view data from a table based on a specific criterion.
(a) form
(b) query
(c) macro
(d) report
34. When entering the field name, what is the maximum number of characters that you can type?
(a) 60
(b) 64
(c) 68
(d) Any number of characters
35. _____ is a database object to view, change, and analyse data in different ways.
(a) Query
(b) Form
(c) Report
(d) None of these
36. Which of the following creates a drop-down list of values to choose from a list?
(a) OLE object
(b) Hyperlink
(c) Memo
(d) Lookup wizard
37. Which field type can store photos?
(a) Hyperlink
(b) OLE
(c) Both (a) and (b)
(d) Access tables cannot store photos
38. Microsoft Access is a(an) _____.
(a) RDBMS
(b) OODBMS
(c) ORDBMS
(d) Network database model
39. What is the full form of SQL?
(a) Structured Query Language
(b) Structured Query List
(c) Simple Query Language
(d) None of these
40. Which of the following SQL statements is correct?
(a) SELECT Username AND Password FROM Users
(b) SELECT Username, Password FROM Users
(c) SELECT Username, Password WHERE Username = 'user1'
(d) All of these

Answers to Multiple-choice Questions

- | | | | | | | | | | |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1 (c) | 2 (d) | 3 (b) | 4 (d) | 5 (b) | 6 (c) | 7 (c) | 8 (a) | 9 (b) | 10 (a) |
| 11 (a) | 12 (b) | 13 (c) | 14 (a) | 15 (b) | 16 (a) | 17 (a) | 18 (b) | 19 (c) | 20 (c) |
| 21 (a) | 22 (a) | 23 (a) | 24 (d) | 25 (a) | 26 (c) | 27 (b) | 28 (b) | 29 (c) | 30 (a) |
| 31 (d) | 32 (c) | 33 (b) | 34 (b) | 35 (a) | 36 (d) | 37 (b) | 38 (a) | 39 (a) | 40 (b) |

Website Designing – HTML

Syllabus Mapping	Unit
Introduction to HTML. Tags and attributes: text formatting, fonts, hypertext links, tables, images, lists, forms, frames, cascading style sheets.	Module II Unit 10

10.1 INTRODUCTION

Hypertext markup language (HTML) is not a programming language, it is a markup language that uses tags to design web pages. It is an interpreted language, a notation for describing document structure (semantic markup) and formatting (presentation markup). Web browsers such as Internet Explorer, Google Chrome, and Mozilla Firefox are used to view HTML documents.

A *web page* is a document or text file commonly written in HTML that is accessible through the Internet using a web browser. A web page is accessed by entering a URL address and may contain text, graphics, and hyperlinks to other web pages.

An HTML document must have a .htm or .html file extension. HTML files can be created with either text editors such as Notepad, Notepad ++ or using HTML editors such as Microsoft FrontPage, Netscape Composer, and Visual Studio.

10.2 STRUCTURE OF HTML PAGE

An HTML page has two distinct parts—the head and the body.

Head This is the header portion. We can use title, script, and metadata tags in this portion. The head is used for text and tags that do not show directly on the page.

Body This is the place where we can enter text, table, graphics, and other work. The body is used for text and tags that are shown directly on the page.

HTML comprises tags (or elements).

The `<!DOCTYPE>` declaration must be the very first thing in your HTML document before the `<html>` tag. It is an instruction to the web browser about what version of HTML the page is written in.

The markup tags provide information about the page content structure. Every tag has opening and closing brackets. Tags may have attributes. They are nested one inside another. Each HTML tag describes a different page content.

There are two types of tags:

Empty tags Tags that do not require an end tag. Example: `
`

Container tags These tags contain both an opening and an end tag.

Examples: `<html>` ... `</html>`, `<head>` ... `</head>`



- The HTML source code should be formatted to increase readability and facilitate debugging.
 - Every block element should start on a new line.
 - Every nested (block) element should be indented.
 - Browsers ignore multiple whitespaces in the page source, so formatting is harmless.
- For performance reasons, formatting can be compromised.

Example 10.1 Create an HTML document.

Solution: Enter the code shown in Fig. 10.1 in Notepad++ and save the file as demo.html. Now run the html page created either by double clicking on it or using the Run option.

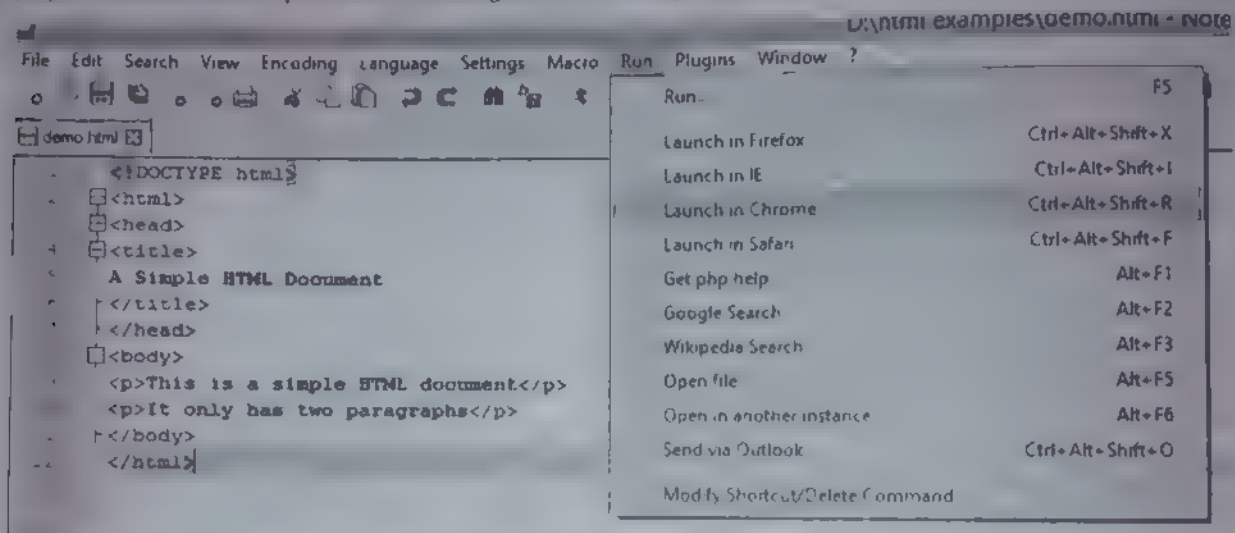


Figure 10.1 Creating an HTML page

This web page will generate the output shown in Fig. 10.2 in the web browser

Displaying the contents of <title> tag

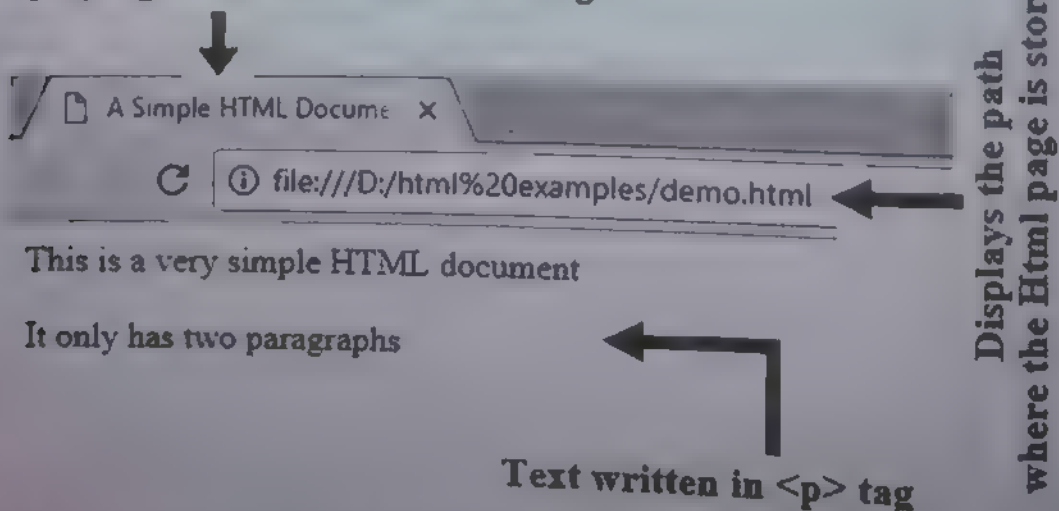


Figure 10.2 Output of HTML code given in Fig. 10.1

10.2.1 Formatting Elements in HTML

The various text formatting elements in HTML are listed in Table 10.1.

Table 10.1 Text formatting elements in HTML

<code><!--></code>	Inserts comments in source code
<code>
</code>	Inserts a single line break
<code><hr></code>	Defines a thematic break in html page
<code></code> <code></code>	Defines bold text
<code><i></code> <code></i></code>	Defines italic text
<code><u></code> <code></u></code>	Defines underline text
<code><big></code> <code></big></code>	Defines bigger text
<code><small></code> <code></small></code>	Defines smaller text
<code></code> <code></code>	Defines the font face, font size and colour of text
<code><h1></code> to <code><h6></code> <code></h1></code> to <code></h6></code>	Headings are defined with <code><h1></code> to <code><h6></code> tags. <code><h1></code> defines the most important heading and <code><h6></code> defines the least important heading
<code><pre></code> <code></pre></code>	Defines preformatted text
<code></code> <code></code>	Defines emphasized text
<code></code> <code></code>	Defines important text
<code><p></code> <code></p></code>	Defines paragraph text
<code><sup></code> <code></sup></code>	Defines superscripted text
<code><sub></code> <code></sub></code>	Defines subscripted text
<code><mark></code> <code></mark></code>	Defines marked/highlighted text



Browsers will display `` as `` and `` as `<i>`. However, there is a difference in the meaning of these tags. `` and `<i>` define bold and italic text, but `` and `` mean that the text is 'important'. Also, `` tag is not supported in HTML5, use CSS instead of `` tag.

Example 10.2 Demonstrate the various text formatting elements in HTML.

Solution: Create an html page say, text_formatting.html containing the source code shown in Fig. 10.3

```

<!DOCTYPE html>
<html>
<body>

<!-- This is a comment -->

<p> font size "3" color "red" < This is some text!> </font> << p>
<p> font face "verdana" color "green" < This is some text!> </font> << p>

<p> To break lines: <br> in a text, <br> use the <br> element << p>

<h1> This is heading 1 </h1>
<h3> This is heading 2 </h3>
<hr>

<h2> HTML <small> Small </small> Formatting </h2>
<h2> HTML <big> big </big> Formatting </h2>
<h2> HTML <mark> Marked</mark> Formatting </h2>
<hr>
<p> This text is normal << p>
<p> <b> This text is bold </b> << p>
<p> <i> This text is italics </i> << p>
<p> <strong> This text is strong</strong> << p>
<p> <em> This text is emphasized </em> << p>
<hr>
<p> This is <sub> subscript</sub> and <sup> superscript</sup> << p>

<pre>
Text in a pre element is displayed in a fixed-width
font, and it preserves both spaces and
line breaks.
</pre>

</body>
</html>

```

Figure 10.3 Source code for Example 10.2

The output is shown in Fig. 10.4.

This is some text!

This is some text!

To break lines
in a text,
use the `br` element.

This is heading 1

This is heading 2

HTML Small Formatting

HTML big Formatting

HTML Marked Formatting

This text is normal

This text is bold

This text is italic

This text is strong

This text is emphasized

This is _{subscript} and ^{superscript}

Text in a `pre` element is displayed in a fixed-width
font, and it preserves both spaces and
line breaks

Figure 10.4 Output of Example 10.2

Paragraph `<p>` Tag

The `<p>` tag has one attribute—`align`, which describes the alignment of text in the paragraph. The list of values that this attribute can take are 'left', 'right', and 'center'.



This `align` attribute is not supported in HTML5.

Example 10.3 Demonstrate the attributes of the `<p>` tag

Solution: Create an html page, say `p_tag.html` containing the source code shown in Fig. 10.5.

```

<!DOCTYPE html>
<html>
<body>

<p align="left">Left justified paragraph</p>
<p align="right">Right justified paragraph</p>
<p align="center">Center justified paragraph</p>

</body>
</html>

```

Figure 10.5 Source code for Example 10.3

The output is shown in Fig. 10.6

Left justified paragraph.

Right justified paragraph.

Center justified paragraph

Figure 10.6 Output of Example 10.3

Image Tag

Image tag is used to display images on an HTML page. Images are linked to the html page using the tag. It creates a holding space for the referenced image. The various attributes of tag are listed in Table 10.2.

Table 10.2 Attributes of tag

	Description
align	Specifies alignment of the image whether top, bottom, right, left, center
src	Points to the path of image file to be displayed
border	Specifies border thickness in pixels
height	Specifies height of image in pixels
width	Specifies width of image in pixels
alt	Specifies alternate text for an image

Example 10.4 Write a code to demonstrate the attributes of the `` tag.

Solution: Create an html page say, image.html containing the source code shown in Fig. 10.7.

```
<!DOCTYPE html>
<html>
<body>
<p><h2> DEMONSTRATING THE IMAGE TAG </h2><br><p>

</body>
</html>
```

Figure 10.7 Source code for Example 10.4

The output is shown in Fig. 10.8.

DEMONSTRATING THE IMAGE TAG



Figure 10.8 Output of example 10.4

Table `<table>` Tag

Tables in HTML are defined row by row. The `<table>` tag creates a table with a border and its value is in pixels. The `<table>` tag itself contains `<tr>` tag to define a new row, `<td>` to define table data or cell, and `<th>` to define table header. The attributes of the `<table>` tag are listed in Table 10.3(a)–(c).

Table 10.3(a) Attributes of table element

	Description
align	Specifies position left, right, center for table
border	Specifies border thickness in pixels (including any cell spacing, default is about 0)
cellspacing	Spacing in pixels between cells, default is about 3
cellpadding	Space in pixels between cell border and table element, default is about 1
width	Width in pixels or percentage of page/frame width

Table 10.3(b) Table row <tr> attributes

align	Specifies position of text left, right, center in table row
bgcolor	Specifies the background colour for table row
valign	Vertically aligns the text in table row

Table 10.3(c) Table cell <td> attributes

colspan	Specifies how many columns this cell occupies
rowspan	Specifies how many rows this cell occupies
bgcolor	Specifies the background color for table cell
valign	Vertically align the text in table cell
width	Width in pixels or percentage of table cell

Example 10.5 Write a code to demonstrate the <table> tag

Solution: Create an html page say, table.html containing the source code as shown in Fig. 10.9.

```
<!DOCTYPE html>
<html>
<body>

<table align="center" width="300" height="200" border="2"
<tr>
  <th>Employee Name</th>
  <th>Department Name</th>
</tr>
<tr align="left" valign="top" bgcolor="red">
  <td>Sonika Vyas</td>
  <td>Accounts</td>
</tr>
<tr align="right" valign="bottom" bgcolor="blue">
  <td>Gaurav Soni</td>
  <td>Technical</td>
</tr>
</table>
```

```

<br>
<TABLE ALIGN "center" WIDTH "300" HEIGHT "200" border "1" >
<TR>
<TD colspan "1" rowspan "2" >a</TD>
<TD colspan "1" rowspan "1" >b</TD>
</TR>
<TR>
<TD colspan "1" rowspan "1" >c</TD>
</TR>
</TABLE>

</body>
</html>

```

Figure 10.9 Source code for Example 10.5

The output is shown in Fig. 10.10.

Employee Name	Department Name
Somika V. Joshi	Accountant

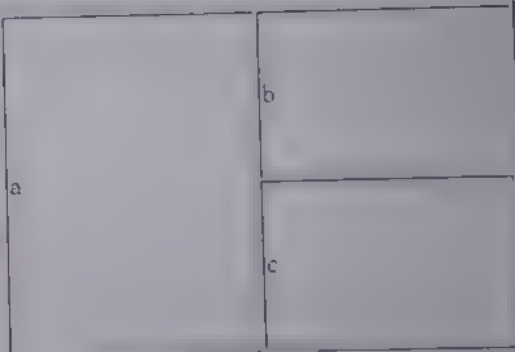


Figure 10.10 Output of Example 10.5

10.3 LISTS IN HTML

HTML provides three ways for specifying lists of information. All lists must contain one or more list elements.

Unordered (bulleted) lists The unordered list is created using the tag. It is a collection of related items that have no special order or sequence. It lists the items using plain bullets. Use the type (it can take 'disc', 'circle', and 'square') attribute to specify the type of bullet; default is 'disc'.

Ordered numbered lists The ordered list is created using the tag. It uses different schemes of numbers to list the items by using the type attribute. The various values the type attribute can have are listed in Table 10.4.

Table 10.4 Attributes of ordered numbered list

<code><ol type = "I"></code>	Default-case numerals
<code><ol type = "I"></code>	Upper-case numerals
<code><ol type = "i"></code>	Lower-case numerals
<code><ol type = "A"></code>	Upper-case letters
<code><ol type = "a"></code>	Lower-case letters

Definition lists Definition lists `<dl>` are those in which entries are listed like in a dictionary or encyclopaedia. The definition list is the ideal way to present a glossary, list of terms, or other name value lists. It uses the `<dt>` and `<dd>` tags inside `<dl>` for a term name and term definition respectively.

Example 10.6 Write a code to demonstrate using lists in HTML.

Solution: Create an html page say, List.html containing the source code as shown in Fig. 10.11.

```

<!DOCTYPE html >
<html >
<body >

<h2 >Unordered HTML List</h2 >

<ul type = "disc">
<li> One </li>
<li> Two </li>
<ul type = "circle">
<li> Three </li>
<li> Four </li>
<ul type = "square">
<li> Five </li>
<li> Six </li>
</ul>
</ul>
</ul>

<h2 >Ordered HTML List</h2 >

<ol type = "I">

```

```

<li>Item one</li>
<li>Item two</li>
<ol type="I">
  <li>Sublist item one</li>
  <li>Sublist item two</li>
</ol>
<ol type="i">
  <li>Sub-sublist item one</li>
  <li>Sub-sublist item two</li>
</ol>
</ol>
</ol>

<h2>Definition HTML List</h2>

<dl>
  <dt><b>HTML</b></dt>
  <dd>This stands for Hyper Text Markup Language</dd>

  <dt><b>HTTP</b></dt>
  <dd>This stands for Hyper Text Transfer Protocol</dd>
</dl>

</body>
</html>

```

Figure 10.11 Source code for Example 10.6

The output is shown in Fig. 10.12

10.4 FRAMES AND FRAMESET IN HTML

The `<frameset>` tag defines a frameset. The `<frameset>` element holds one or more `<frame>` elements. Each `<frame>` element can hold a separate document. The `<frameset>` element specifies how many columns or rows there will be in the frameset, and how much percentage pixels of space will each of them occupy. Table 10.5 lists the attributes of the `<frame>` tag.

Unordered HTML List

- One
- Two
 - Three
 - Four
 - Five
 - Six

Ordered HTML List

- 1 Item one
- 2 Item two
 - I Sublist item one
 - II Sublist item two
 - Sub-sublist item one
 - Sub-sublist item two

Definition HTML List

- HTML**
This stands for Hyper Text Markup Language
- HTTP**
This stands for Hyper Text Transfer Protocol

Figure 10.12 Output of Example 10.6



Frames are not supported in HTML5.

Table 10.5 Attributes of <frame> tag

src	Specifies the URL of the document to be shown in a frame
marginheight	Specifies the top and bottom margins of a frame in pixels
marginwidth	Specifies the left and right margins of a frame in pixels
name	Specifies the name of a frame
noresize	Specifies that a frame is not resizable
scrolling	Specifies whether or not to display scrollbars in a frame. It can take yes, no, auto values
frameborder	Specifies whether border is needed or not. It can take either 0 or 1 value

Example 10.7 Write a code to demonstrate the <frame> and <frameset> tags.

Solution: Create the html pages shown in Figs. 10.13(a)–(d).

```
<!DOCTYPE html>
<html>

  <head>
    <title>HTML Frames</title>
  </head>

  <frameset rows="10%,80%,10%">
    <frame name="top" src="top.html" />
    <frame name="main" src="main.html" />
    <frame name="bottom" src="bottom.html" />
  </frameset>

</html>
```



```
<!DOCTYPE html>
<html>
```

```
<body>
<h3>THIS IS THE TOP FRAME</h3>
</body>
</html>
```

(b)

```
<!DOCTYPE html>
<html>
```

```
<body>
<h3>THIS IS THE MAIN FRAME</h3>
</body>
</html>
```

(c)

```
<!DOCTYPE html>
<html>
```

```
<body>
<h3>THIS IS THE BOTTOM FRAME</h3>
</body>
</html>
```

(d)

Figure 10.13 Codes for Example 10.7 (a) frameset.html (b) top.html (c) main.html (d) bottom.html

The output is shown in fig. 10.14

THIS IS THE TOP FRAME

THIS IS THE MAIN FRAME

THIS IS THE BOTTOM FRAME

Figure 10.14 Output of Example 10.7

10.5 OTHER TAGS

In this section, we will discuss the input, form, and link tags of HTML.

10.5.1 Input Tag

The `<input>` tag is used to select user information. Its elements are used within a `<form>` element to declare input controls that allow users to input data. An input field can vary in many ways, depending on the type attribute. The attributes of the `<input>` tag are listed in Table 10.6.

Table 10.6 Attributes of `<input>` tag

Autocomplete	On, Off	Specifies whether an <code><input></code> element should have autocomplete enabled or not
Autofocus	Autofocus	Specifies that an <code><input></code> element should automatically get focus when the page loads
Checked	Checked	Specifies that an <code><input></code> element should be pre-selected when the page loads (for <code>type="checkbox"</code> or <code>type="radio"</code>)
Disabled	Disabled	Specifies that an <code><input></code> element should be disabled
Max	Number	Specifies the maximum value for an <code><input></code> element
Maxlength	Number	Specifies the maximum number of characters allowed in an <code><input></code> element
Min	Number	Specifies the minimum value for an <code><input></code> element
Placeholder	Text	Specifies a short hint that describes the expected value of an <code><input></code> element
Readonly	Readonly	Specifies that an input field is read-only
Required	Required	Specifies that an input field must be filled out before submitting the form
Type	Button, checkbox, email, hidden, password, number, radio, submit, text, reset	Specifies the type <code><input></code> element to display

10.5.2 Form Tag

The `<form>` tag is used to create an html form for user input. An HTML form is used to pass data to a server. It can contain one or more of the following form elements: `<input>`, `<button>`, `<select>`, `<option>`. It is a block-level element, and browsers create a line break before and after a form. Table 10.7 lists the attributes this tag can take.

Table 10.7 Attributes of <form> tag

Action	Specifies where to send the form-data when a form is submitted Value can be absolute or relative url
Enctype	Specifies how the form-data should be encoded when submitting it to the server (only for method="post")
Name	Specifies name of the form
Method	Specifies how to send form-data (the form-data is sent to the page specified in the action attribute)
Target	Specifies the target window or frame where the result of the script will be displayed It takes values such as _blank, self, parent

get Method

- This method appends the form-data to the URL in name/value pairs.
- There is a limit to how much data you can place in a URL (varies between browsers), therefore, you cannot be sure that all the form-data will be correctly transferred.
- Never use the get method to pass sensitive information (password or other sensitive information will be visible in the browser's address bar).

post Method

- This method sends the form-data as an HTTP post transaction.
- Form submissions with the post method cannot be bookmarked
- The post method is more robust and secure than 'get', and it does not have size limitations

Example 10.8 Write a code to demonstrate the <form> tag.

Solution: Create an html page say, form.html containing the source code shown in Fig. 10.15.

```
<!DOCTYPE html>
<html>

  <head>
    <title>To Get User Name</title>
  </head>

  <body>
    <form>
      First name <input type="text" name="first_name">
      <br>
      Last name <input type="text" name="last_name">
    </form>
  </body>
</html>
```

```

        <br><br>
        <input type="submit" value="Submit">
    </form>
</body>

</html>

```

Figure 10.15 Source code for Example 10.8

The output is shown in Fig. 10.16.

First name:

Last name:

Figure 10.16 Output of Example 10.8

10.5.3 Link Tag

The `<link>` tag defines the relationship between a document and an external resource. The `<link>` tag is most used to link to style sheets. It is an empty element, it contains only attributes. This element goes only in the head section, but it can appear any number of times. The attributes include href, rel, target, type, etc.

10.6 CASCADING STYLE SHEETS

Styles define how to display HTML elements. Styles are normally stored in style sheets. Cascading style sheets use the `<style>` tag for styling. The syntax is as follows:

```
selector {property: value}
```

The selector is normally the HTML element/tag. The property is the attribute you wish to change and each property can take a value.

10.6.1 Methods of Specifying CSS

The following are the three methods to specify CSS.

Tag Modifier

They specify a similar styling for the HTML tags used in the entire page.

Example 10.9 Write a code to demonstrate the tag modifier CSS.

Solution: Create an HTML page say, css_demo.html containing the source code as shown in Fig. 10.17

```
<!DOCTYPE html >
<html>
<head>
<style>
body {
    background-color: lightpink;
}

h1 {
    color: black;
    text-align: center;
}

p {
    font-family: verdana;
    font-size: 20px;
}
</style>
</head>
<body>

<h1>Learning CSS</h1>
<p>This is a paragraph</p>

</body>
</html>
```

Figure 10.17 Source code for Example 10.9

The output is shown in Fig. 10.18.

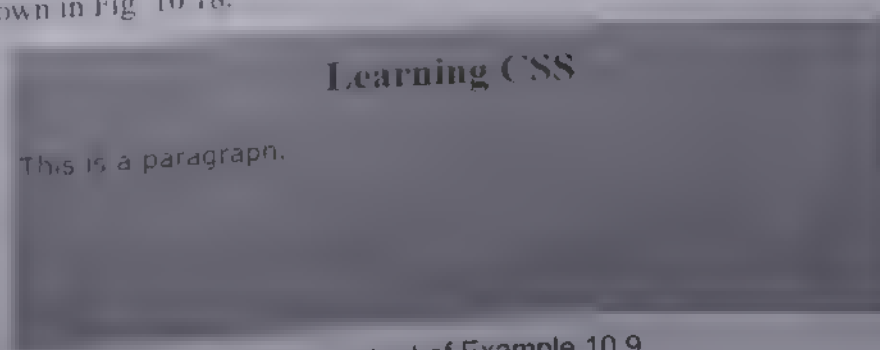


Figure 10.18 Output of Example 10.9

Class Modifier

With the class selector, you can define different styles for the same type of HTML element.

Example 10.10 Write a code to demonstrate the class modifier CSS.

Solution: Create an html page say, class_css.html containing the source code as shown in Fig. 10.19.

```
<!DOCTYPE html>
<html>
<head>
<style>

        center {text-align: center;
                color: red;}
        bg_col {background-color: lightblue;}

</style>
</head>
<body class="bg_col">
<h1 class="center">
        This heading will be center-aligned
</h1>

</body>
</html>
```

Figure 10.19 Source code for Example 10.10

The output is shown in Fig. 10.20.

id Selector

With the id selector you can define the same style for different HTML elements. The id selector uses the id attribute of the HTML element, and is defined with a "#". The style rule below will be applied to the element with id="para1":

Example 10.11 Write a code to demonstrate the id selector CSS.

Solution: Create an HTML page say, id_css.html containing the source code as shown in Fig. 10.21.

```
<!DOCTYPE html>
<html>
<head>
<style>
```

This heading will be center aligned

Figure 10.20 Output of Example 10.10

```
#green {color: green;}

#para1
{
text-align:center;
color:red;
}

#bg_col
{ background-color:cyan;
}

< style >
< head >

< body id "bg_col"
< h1 id "green" ID SELECTOR CSS > h1
< p id "para1" > This paragraph will be center aligned < /p

< / body >
< /html >
```

Figure 10.21 Source code for Example 10.11

The output is shown in Fig. 10.22

ID SELECTOR CSS

www.it-ebooks.info

Figure 10.22 Output of Example 10.11

10.6.3 Ways of Inserting a Style Sheet

There are three ways of inserting a style sheet: internal, inline, and external.

Internal Style

An internal style sheet should be used when a single document has a unique style. You define internal styles in the head section by using the < style > tag.

Example 10.12 Write a code to demonstrate internal styles.

Solution: Create an html page say, int_css.html containing the source code as shown in Fig. 10.23

```
<!DOCTYPE html>
<html>
  <head>
    <style type="text/css">
      h1 {color:blue; text-align:center;}
      p {margin-left: 20px; text-decoration: line-through;}
      body {background-color:lime;}
    </style>
  </head>
  <body>
    <h1>WE ARE LEARNING CSS</h1>
    <p>THIS IS INTERNAL STYLING</p>
  </body>
</html>
```

Figure 10.23 Source code for Example 10.12

The output is shown in Fig. 10.24.



Figure 10.24 Output of Example 10.12

Inline Style

An inline style may be used to apply a unique style for a single element.

Example 10.13 Write a code to demonstrate inline styles.

Solution: Create an html page say, style.html containing the source code as shown in Fig. 10.25.

```
<!DOCTYPE html>
<html>
  <body>

    <h1 style="color:blue;margin-left:30px;">This is a heading</h1>
    <p style="text-decoration: overline; text-transform: uppercase;">This is a paragraph.</p>

  </body>
</html>
```

Figure 10.25 Source code for Example 10.13

The output is shown in Fig. 10.26.

External Style

An external style sheet can be written in any text editor. The file should not contain any HTML tags. The style sheet file must be saved with a .css extension. With an external style sheet, you can change the look of an entire website by changing just one file. Each page must include a reference to the external style sheet file inside the <link> element.

Example 10.14 Write a code to demonstrate external style sheet.

Solution: Create an html page which contains a reference to an external css file (Fig. 10.27).

```
<!DOCTYPE html>
<html>
<head>
<link rel="stylesheet" type="text/css" href "ext_style.css" >
</head>
<body>

<h1>This is a heading</h1>
<p>This is a paragraph.</p>

</body>
</html>
```

(a)

```
body {
background-color: yellow;
}

h1 {
color: grey;
margin-left: 50px;
text-decoration: underline;
}
```

(b)

Figure 10.27 Source code for Example 10.14
(a) ex_css.html (b) ext_style.css

The output is shown in Fig. 10.28.

This is a heading

THIS IS A PARAGRAPH.

Figure 10.26 Output of Example 10.13

This is a heading

This is a paragraph

Figure 10.28 Output of Example 10.14

10.6.4 Cascading Order (Priority) of Styling

- Browser default (highest priority)
- External style sheet
- Internal style sheet (in the head section)
- Inline style (inside an HTML element) (lowest priority)

Example 10.15 Write a code to demonstrate the CSS font properties.

Solution: Create an html page say, `css_fonts.html` containing the source code as shown in Fig. 10.29.

```
<DOCTYPE html>
<html>
  <head>
    <style>
      p.normal {
        font-weight: normal,
        font-family: 20px Verdana;
      }

      p.light {
        font-weight: lighter,
        font-family: "Times New Roman", Times, serif;
      }

      p.thick {
        font-weight: bold,
        font-style: oblique;
        font-size: 30px;
      }
    </style>
  </head>
  <body>
    <p class="normal">This is a paragraph.</p>
    <p class="light">This is a paragraph.</p>
    <p class="thick">This is a paragraph.</p>
  </body>
</html>
```

Figure 10.29 Source code for Example 10.15

The output is shown in Fig. 10.30.

Example 10.16 Create the registration form using HTML, create the CSS file, and implement the CSS with HTML.

Solution: The source codes are shown in Figs. 10.31(a) and (b).

This is a paragraph.

This is a paragraph.

This is a paragraph.

Figure 10.30 Output of Example 10.15

```
<!DOCTYPE html>
<html>
<head>
<title>this is registration form</title>
<link rel="stylesheet" href="styles.css">
</head>
<body>
<table cellpadding="2" width="20%" align="center"
cellspacing="2">
<tr>
<td colspan="2">
<center><font size="4"><b>Student Registration Form</b></center></td></tr>
<tr>
<td>Name</td>
<td><input type="text" name="textnames" id="textname" size="30"></td></tr>
<tr>
<td>Gender</td>
<td><input type="radio" name="sex" value="male" size="10">Male
<input type="radio" name="sex" value="female" size="10">Female</td></tr>
<tr>
<td>Course</td>
<td><select name="Course">
<option value="-1" selected></option>
<option value="B.Tech">B.TECH</option>
<option value="MCA">MCA</option>
<option value="MBA">MBA</option>
<option value="BCA">BCA</option>
</select></td></tr>
<tr>
<td>State</td>
<td><select Name="State">
```

```

    <option value="1" selected>select </option>
    <option value="New Delhi">NEW DELHI</option>
    <option value="Mumbai">MUMBAI</option>
    <option value="Goa">GOA</option>
    <option value="Bihar">BIHAR</option>
  </select> <td> <tr>
<tr>
<td>PinCode</td>
<td><input type="text" name="pincode" id="pincode" size="30"></td></tr>
<tr>
<td>EmailId</td>
<td><input type="text" name="emailid" id="emailid" size="30"></td></tr>
<tr>
<td>DOB</td>
<td><input type="text" name="dob" id="dob" size="30"></td>
</tr>
<tr>
<td>MobileNo</td>
<td><input type="text" name="mobileno" id="mobileno" size="30"></td>
</tr>
<tr>
<td><input type="reset"></td>
<td colspan="2"><input type="submit" value="Submit Form"></td>
</tr>
</table></form>
</body>
</html>

```

(a)

```

input[type="text"] {
background-color: yellow;
color: purple;}
body{
background-color: lightblue;}
form{
background-color: green;}

```

(b)

Figure 10.31 Source code for Example 10.16 (a) Student.html (b) Styles.css

The output is shown in Fig. 10.32.

Figure 10.32 Output of Example 10.16

Summary

- An HTML document must have a .htm or .html file extension.
- An HTML page has two distinct parts—the head and the body.
- Headings are defined with <h1> to <h6> tags. <h1> defines the most important heading and <h6> defines the least important heading.
-
 tag defines a single line break.
- The align attribute of <p> tag describes the alignment of text inside the paragraph. It can have 'left', 'right', and 'center' values.
- The <table> tag itself contains <tr> tag to define a new row, <td> to define table data or cell and <th> to define table header.
- Each <frame> element can hold a separate document.
- Frameborder attribute of frame tag specifies whether border is needed or not.
- HTML provides three ways for specifying lists of information: unordered lists, ordered lists, and definition lists.
- Get method appends the form-data to the URL in name/value pairs.
- Post method sends the form-data as an HTTP post transaction.
- Form tag can contain one or more of the following form elements: <input>, <button>, <select>, <option>, etc. It is a block-level element, and browsers create a line break before and after a form.
- The style sheet file must be saved with a .css extension.
- Tag modifier specifies the similar styling for html tags used in the entire html page.
- An inline style may be used to apply a unique style for a single element.

Glossary

Definition lists Those in which entries are listed like in a dictionary or encyclopaedia

External style sheet This is used to change the look of an entire website by changing just one file

Frameset This defines a frameset. The frameset element holds one or more <frame> elements

Form It is used to create an HTML form for user input. An HTML form is used to pass data to a server

HTML A markup language that uses tags to design web pages

Image tag It is used to display images on an HTML page

Internal style sheet This is used when a single document has a unique style

Link It defines the relationship between a document and an external resource

Markup tags These provide information about the page content structure

Styles They define how to display HTML elements

Table tag It creates a table with a border and its value is in pixels

Web page A document or text file commonly written in HTML that is accessible through the Internet using a web browser

Multiple-choice Questions

- What type of language is HTML?
 - Scripting language
 - Markup language
 - Programming language
 - Network protocol
- HTML uses _____.
 - user-defined tags
 - pre-specified tags
 - fixed tags defined by the language
 - tags only for linking
- The fundamental HTML block is known as _____.
 - HTML body
 - HTML tag
 - HTML attribute
 - HTML element
- Apart from the `` tag, which other tag makes text bold?
 - `<fat>`
 - ``
 - `<black>`
 - `<emp>`
- HTML web pages can be read and rendered by _____.
 - compiler
 - server
 - web browser
 - interpreter
- Which of the following is not a browser?
 - Microsoft Bing
 - Netscape Navigator
 - Mozilla Firefox
 - Opera
- Which HTML tag produces the biggest heading?
 - `<h7>`
 - `<h9>`
 - `<h4>`
 - `<h1>`
- HTML tags are surrounded by which type of brackets?
 - Curly
 - Squart
 - Round
 - Angle
- Tags and text that are not directly displayed on the page are written in the _____ section.
 - `<head>`
 - `<title>`
 - `<body>`
 - `<html>`
- The common element which describes the web page is the _____.
 - heading
 - paragraph
 - list
 - all of these
- From which tag do descriptive lists start?
 - `<LL>`
 - `<DD>`
 - `<DL>`
 - `<DS>`
- The attribute of `<form>` tag is _____.
 - method
 - action
 - both (a) and (b)
 - none of these
- Markup tags tell the web browser how to _____.
 - organize the page
 - display the page
 - display a message box on the page
 - none of these
- What are empty elements and are they valid?
 - No, there is no such term
 - Empty elements are elements with no data
 - No, it is not valid to use empty elements
 - None of these
- Which of the following attributes of text box control allows the user to limit the maximum character?
 - Size
 - Len
 - Maxlength
 - All of these

16. Web pages start with which of the following tags?
 - (a) <Body>
 - (b) <Title>
 - (c) <HTML>
 - (d) <Form>
17. Which of the following is a container?
 - (a) <SELECT>
 - (b) <Value>
 - (c) <INPUT>
 - (d) <BODY>
18. <DT> tag is designed to fit a single line of our web page but <DD> tag will accept a _____.
 - (a) line of text
 - (b) full paragraph
 - (c) word
 - (d) request
19. The correct HTML code to left align the content inside a table cell is _____.
 - (a) <td left>
 - (b) <td raligh = "left">
 - (c) <td align = "left">
 - (d) <td leftalign>
20. How can you open a link in a new browser window?
 - (a)
 - (b)
 - (c)
 - (d)
21. Which of the following tags is used to create a numbered list?
 - (a)
 - (b)
 - (c) and
 - (d) None of these
22. The tag used to create a hypertext relationship between the current document and another URL is _____.
 - (a) <ISINDEX>
 - (b) <A>
 - (c) <LINK>
 - (d) none of these
23. The text inside the <TEXT AREA> tag works like _____.
 - (a) <P> formatted text
 - (b) <T> formatted text
 - (c) <PRE> formatted text
 - (d) none of these
24. What is the correct HTML code for adding a background color?
 - (a) <background>yellow<Background>
 - (b) <body color = "yellow">
 - (c) <body bg color = "yellow">
 - (d) <body bg ="yellow">
25. The main container for <TR>, <TD>, and <TH> is _____.
 - (a) <TABLE>
 - (b) <GROUP>
 - (c) <DATA>
 - (d) All of these
26. How can you make an e-mail link?
 - (a) <mail href+"xxx@y.com">
 - (b)
 - (c)
 - (d) Both (b) and (c)
27. The web standard allows programmers on many different computer platforms to disperse format and display the information server. These programs are called _____.
 - (a) web browsers
 - (b) HTML
 - (c) Internet Explorer
 - (d) none of these
28. What is the correct HTML code for referring to an external style sheet?
 - (a) <link rel="stylesheet" type="text/css" href="mystyle.css">
 - (b) <style src="mystyle.css">
 - (c) <stylesheet>mystyle.css</stylesheet>
 - (d) None of these
29. Where in an HTML document is the correct place to refer to an external style sheet?
 - (a) At the end of the document
 - (b) In the <body> section
 - (c) In the <head> section
 - (d) Both (a) and (b)
30. Which HTML tag is used to define an internal style sheet?
 - (a) <css>
 - (b) <style>
 - (c) <script>
 - (d) None of these
31. Which HTML attribute is used to define inline styles?
 - (a) Style
 - (b) Class
 - (c) Styles
 - (d) Font
32. Which is the correct CSS syntax?
 - (a) body:color=black;
 - (b) body {color: black;}
 - (c) {body:color=black;}
 - (d) {body;color:black;}

33. How do you insert a comment in a CSS file?
 (a) `// this is a comment //`
 (b) `// this is a comment`
 (c) `* this is a comment`
 (d) `/* this is a comment */`
34. How do you add a background color for all `<h1>` elements?
 (a) `all.h1 {background-color:#FFFFFF;}`
 (b) `h1 {background-color:#FFFFFF;}`
 (c) `h1.all {background-color:#FFFFFF;}`
 (d) None of these
35. Which CSS property controls the text size?
 (a) Font-style (c) Text-size
 (b) Font-size (d) Text-style
36. What is the correct CSS syntax for making all the `<p>` elements bold?
 (a) `p {font-weight:bold;}`
 (b) `p {text-size:bold;}`
 (c) `<p style="text-size:bold;">`
 (d) `<p style="font-size:bold;">`
37. Which property is used to change the left margin of an element?
 (a) margin-left (c) padding-left
 (b) indent (d) None of these
38. How do you select an element with id "demo"?
 (a) `demo` (c) `*demo`
 (b) `#demo` (d) `.demo`
39. How do you select elements with class name "test"?
 (a) `.test` (c) `#test`
 (b) `test` (d) `*test`
40. What is the default value of the position property?
 (a) Relative (c) Fixed
 (b) Static (d) Absolute

Answers to Multiple-choice Questions

1. (b) 2. (c) 3. (b) 4. (b) 5. (c) 6. (a) 7. (d) 8. (d) 9. (a) 10. (d)
 11. (c) 12. (c) 13. (b) 14. (b) 15. (c) 16. (c) 17. (a) 18. (b) 19. (c) 20. (b)
 21. (c) 22. (c) 23. (c) 24. (c) 25. (a) 26. (b) 27. (a) 28. (a) 29. (c) 30. (b)
 31. (a) 32. (b) 33. (d) 34. (b) 35. (b) 36. (a) 37. (a) 38. (b) 39. (a) 40. (b)